



# Toward a more humane genetics education: Learning about the social and quantitative complexities of human genetic variation research could reduce racial bias in adolescent and adult populations

Brian M. Donovan<sup>1</sup> | Rob Semmens<sup>2</sup> | Phillip Keck<sup>3</sup> | Elizabeth Brimhall<sup>4</sup> | K. C. Busch<sup>5</sup> | Monica Weindling<sup>1</sup> | Alex Duncan<sup>1</sup> | Molly Stuhlsatz<sup>1</sup> | Zoë Buck Bracey<sup>1</sup> | Mark Bloom<sup>1</sup> | Susan Kowalski<sup>1</sup> | Brae Salazar<sup>1</sup>

<sup>1</sup>Biological Sciences Curriculum Study (BSCS) Science Learning, Colorado Springs, Colorado

<sup>2</sup>Department of Systems Engineering, Naval Post Graduate School, Monterey, California

<sup>3</sup>The Live Oak School, San Francisco, California

<sup>4</sup>Palo Alto Unified School District, Palo Alto, California

<sup>5</sup>College of Education, North Carolina State University, Raleigh, North Carolina

## Correspondence

Brian M. Donovan, Biological Sciences Curriculum Study (BSCS) Science Learning, 5415 Mark Dabbling Boulevard, Colorado Springs, CO 80918.  
Email: bdonovan@bscs.org

## Funding information

National Science Foundation, Grant/Award Number: 1660985

## Abstract

When people are exposed to information that leads them to overestimate the actual amount of genetic difference between racial groups, it can augment their racial biases. However, there is apparently no research that explores if the reverse is possible. Does teaching adolescents scientifically accurate information about genetic variation within and between US census races reduce their racial biases? We randomized 8<sup>th</sup> and 9<sup>th</sup> grade students ( $n = 166$ ) into separate classrooms to learn for an entire week either about the topics of (a) human genetic variation or (b) climate variation. In a cross-over randomized trial with clustering, we demonstrate that when students learn about genetic variation within and between racial groups it significantly changes their perceptions of human genetic variation, thereby causing a significant decrease in their scores on instruments assessing cognitive forms of prejudice. We then

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2019 The Authors. *Science Education* Published by Wiley Periodicals, Inc.

replicate these findings in two computer-based randomized controlled trials, one with adults ( $n = 176$ ) and another with biology students ( $n = 721$ , 9<sup>th</sup>–12<sup>th</sup> graders). These results indicate that teaching about human variation in the domain of genetics has potentially powerful effects on social cognition during adolescence. In turn, we argue that learning about the social and quantitative complexities of human genetic variation research could prepare students to become informed participants in a society where human genetics is invoked as a rationale in sociopolitical debates.

#### KEYWORDS

analogical transfer, genetic essentialism, genetics education, racial prejudice, social cognition

## 1 | INTRODUCTION

Throughout history, the science of genetics has been used to justify policies that helped one race by harming another (Jackson & Depew, 2017). Even today the belief that races are genetically dissimilar is used to justify interethnic hostility (Kimel, Huesmann, Kunst, & Halperin, 2016), prejudice (Dar-Nimrod & Heine, 2011), segregative behavior (Williams & Eberhardt, 2008), and discriminatory policies (Soylu, Estrada-Villalta, & Adams, 2017). Evidence suggests that the biology curriculum plays a role in the perpetuation of this problem (Donovan, 2015b). For example, early 20<sup>th</sup> century biology textbooks taught students about the biological superiority of the White race (Donovan, 2015b; Morning, 2011; Willinsky, 1998). Although such claims are absent from contemporary texts (Donovan, 2015b; Morning, 2011), randomized controlled trials (RCTs) have demonstrated a cause–effect relationship between the treatment of race in the biology curriculum and the development of racial biases (Donovan, 2014, 2016, 2017). When students learn about the prevalence of particular genetic diseases in specific racial groups during middle or high school biology classes, it can unintentionally lead youth to perceive more genetic variation between races than actually exists (Donovan, 2017) and thus infer that racial groups differ in intelligence for genetic reasons (Donovan, 2014, 2016, 2017). In turn, this learning appears to affect students' support for policies that redress racial inequality by influencing how students explain racial disparities (Donovan, 2016, 2017).

A biology curriculum that perpetuates racial bias by unintentionally increasing inaccurate beliefs about racial difference is inhumane because it harms those who suffer from racial discrimination. In this study, we answer the question: Does learning scientifically accurate information about genetic variation within and between US census racial groups cause a significant reduction in racial bias? We demonstrate that teaching scientifically accurate information about genetic variation within and between the US census races can reduce racial bias by undermining the belief that racial groups are discrete. These findings establish *preliminary* proof of concept for the hypothesis that human genetics education can be designed to create a more humane society by reducing an individual's racial biases. Although the stereotypes we attempt to reduce through this study are focused on the particular demographic categories associated with America's legacies of racism and genocide, we suggest that our work is relevant in other nations.

Over the last 300 years the oppression and subjugation of various ethnic groups around the world has been rationalized through the idea that inequality is a natural product of human biological difference (Doron, 2012; Hudson, 1996; Mallon, 2013; Morning, 2011; Smedley & Smedley, 2005; Wolf, Kahn, Roseberry, & Wallerstein, 1994). Particularly in Europe and in those continents colonized by Europeans, science education has also been used to lend false credence to the belief that Western Europeans are biologically superior to other groups (Morning,

2011; Willinsky, 1998). Unfortunately, this belief is still apparent in education today as studies estimate that 3–62% of biology teachers in European, South-American, African, or Middle-Eastern countries agree that, “Ethnic groups are genetically different and that is why some are superior to others” (e.g., 3% in France; 18% in Senegal; 34% in Poland; and 62% in Lebanon; Castéra & Clément, 2014). We, therefore, contend that human genetics education is not socially neutral in any society where genetics is used to rationalize prejudice and that is why the humane genetics education hypothesis we advance may be relevant to genetics education outside of the United States.

## 2 | THE HUMANE GENETICS EDUCATION HYPOTHESIS

The purpose of a humane genetics education is to reduce the prevalence of racial bias by changing the way that students perceive human genetic variation. Such a hypothesis raises three important questions. First, how much biological variation actually exists between and within racially defined groups? Second, why would teaching students this genetic content lead to lower levels of racial bias? After all, the belief that races differ genetically is significantly associated with the belief that racial inequality is not worthy of redress because it is a natural and unchangeable product of genes (Brueckner, Morning, & Nelson, 2005; Donovan, 2015a; Jayaratne et al., 2006; Morning, 2011). Thus, it is reasonable to predict that teaching about human genetic difference will increase racial bias. Indeed, one might point to the long history of genetically justified racial bias in education (e.g., Jensen, 1969) as a reason for why it is not wise to discuss genetic differences between races in school biology. Yet, history is filled with scientists who challenged racism by pointing out the genetic flaws in racist beliefs (e.g., Beckwith, 2009; Feldman & Lewontin, 1975; Gould, 1996; Graves, 2015; Lewontin, 1972; Livingstone & Dobzhansky, 1962). Those scientists have argued that accurate understandings of genetic variation undermine the apparent validity of prejudiced beliefs (Jackson & Depew, 2017).

Prejudiced beliefs based on genetic thinking share a set of assumptions that are scientifically flawed (Jackson & Depew, 2017). First is the premise that people of the same race are genetically uniform. Second is the premise that people of disparate races are categorically different. When these assumptions are combined with the belief that biologically influenced abilities are immutable, people will then argue that it is pointless to intervene socially to reduce racial inequality, because race biologically determines ability (Lewontin, 1996).

Such arguments may seem anachronistic in the 21<sup>st</sup> century. However, there are many White supremacist websites that reach large audiences around the world (Jacobs, 2015), such as *The Daily Stormer* (2016) or “White Pride For Kids” (n.d.). On such sites, one can read arguments such as, “the Jews” are “a separate race, with biological drives and behavior patterns which come into direct conflict with the goals and values of the White race” (*The Daily Stormer*, 2016). Knowing why these arguments are flawed is the content of a humane genetics education.

### 2.1 | The content of a humane genetics education

Any two humans share 99.9% of their DNA, which means that 0.1% of human DNA varies between individuals. Studies find that, on average, 4.3% of genetic variability in humans (4.3% of the 0.1% of the variable portion of human DNA) occurs between the continental populations commonly associated with US census racial groups (i.e., Africa, Asia, Pacific Islands, and The Americas, Europe). In contrast, 95.7% of human genetic variation (95.7% of the 0.1% of variable portion of human DNA) occurs between individuals within those same groups (Rosenberg, 2011; Rosenberg et al., 2002). Thus, if we randomly pick two individuals from two different continental groups and compare them with two randomly picked individuals from the same group, we can expect that the former will be 4.3% more different from one another than the latter (Donovan, 2015a). These findings undermine the idea of intraracial uniformity because they show that people of the same group are different in their variable DNA. They also show that racial groups are not discrete.

In fact, this same pattern is found repeatedly in studies of human variation (Boas, 1911; Relethford, 2002). There is more variability in skull shape, facial structure, and blood types within racially defined populations (e.g., 89% for craniometric traits and 86% for blood types) than there is between them (Relethford, 2002).

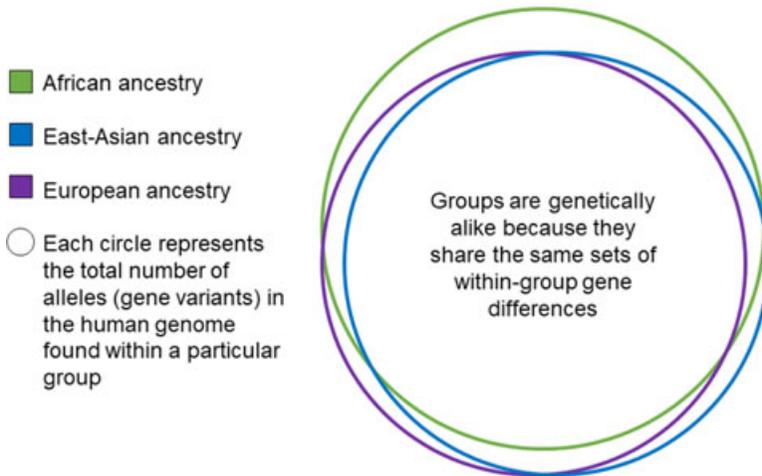
Population genetic theory predicts that this same pattern will hold for any human trait that is not under different selection pressures in disparate populations (Edge & Rosenberg, 2015).

One known exception to this pattern is skin color, which varies more between populations than within populations (Relethford, 2002). Skin color changes continuously with distance from the equator (Relethford, 1997, 2002), because it is an adaptation to latitudinal differences in ultraviolet radiation (Jablonski & Chaplin, 2010). Therefore, there is no place on the planet where one can draw a line separating “dark-skinned” populations from “light-skinned” populations. Also, the gene variants associated with dark and light skin color are found in every major human population (Crawford et al., 2017). This means that the alleles associated with “light skin” are found in “dark-skinned” populations and vice versa (Crawford et al., 2017). The variants associated with dark or light skin are simply more prevalent in certain populations (Crawford et al., 2017). Both the continuous variation in skin color and the fact that ancestry groups share the genes associated with it undermine the idea that “Blacks” and “Whites” are discrete. Furthermore, because patterns of skin color variation do not match patterns of variation in other traits (i.e., blood types or skull shape), skin color cannot be used to make valid inferences about the geographic variation of other biological traits (Relethford, 2002). Thus, even if human behavior and cognition were genetically determined (and they are not, see Beckwith & Pierce, 2018), knowing a stranger’s skin color would still not permit an accurate prediction about their abilities.

As uniformity and discreteness beliefs about race are biologically flawed, it should come as no surprise that there has never been any agreement within the biological or anthropological sciences about whether human races are biologically real (Jackson & Depew, 2017; Morning, 2011). Even today there is no scientific consensus that race is biologically real (Foster, 2009; Kaplan & Winther, 2013a, 2013b; Weiss, 2008; Yudell, Roberts, DeSalle, & Tishkoff, 2016). There is, however, a consensus that population genetic data does not support racist claims (Coop, Eisen, Nielsen, Przeworski, & Rosenberg, 2014). For instance, physical anthropologists and population geneticists have recently argued that genetic data (a) refutes the notion that races are biologically real (Hunley, Cabana, & Long, 2016), (b) cannot be used to support a biological theory of race without many additional philosophical arguments (Rosenberg & Edge, in press), (c) refutes the idea of racially pure populations (Reich, 2018), and (d) refutes White supremacist beliefs (“ASHG Denounces Attempts to Link Genetics and Racial Supremacy”, 2018). Racial categories are an artifact of culture, history, and beliefs, but not biology (Markus & Moya, 2011).

This does not mean, however, that there are absolutely no biological differences between the socially created racial groups used in the United States. If this were true, then all variation would occur within populations and there would be no variation between geographic populations. In other words, *Homo sapiens* would be a single population. Instead, geographic populations of humans do differ. Yet, they do not differ in the ways that many people might think. Human groups differ in the proportion of people who have certain gene variants (Rosenberg, 2011). Although some gene variants are unique to a single group (7.53%; Rosenberg, 2011), on average none of those unique variants are possessed by >1.65% of any population (Rosenberg, 2011). Furthermore, on average, the amount of genetic difference between geographic groups of humans is about seven times less than the genetic differences between populations of chimpanzees (Becquet, Patterson, Stone, Przeworski, & Reich, 2007) and about 14 times less than the genetic differences between populations of White-tailed deer (Graves, 2015). The relatively small genetic differentiation between human groups is a product of the finding that the majority of variants within the human genome are found in two or more continental groups of humans (Rosenberg, 2011). Therefore, genomic findings about human ancestry groups clearly demonstrate that the culturally specific racial groupings in the US census are genetically alike. Furthermore, US races are genetically alike because they share similar sets of within-group differences in their variable DNA (Figure 1).

Nevertheless, the small proportional differences found when comparing the genomes of different US racial categories are often used to claim that racial inequality in the United States is immutable because of genetics. Jensen (1969) was well known for making this kind of claim. To understand why Jensen’s claim and more modern versions of it are wrong (e.g., Wade, 2014) a proper understanding of genetic variation within and between populations is required. Specifically, one needs to understand that the causes of within-group variation in a trait can be different from the causes



**FIGURE 1** Genetic differences within and between groups. Adapted from Rosenberg (2011). Data can be found on p. 665 [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

of between-group variation in that same trait. That is, even when trait differences between individuals within a population are entirely inherited, differences between populations can still be caused entirely by environmental factors (Feldman & Lewontin, 1975). Indeed, if you estimate the heritability of skin color among White New Yorkers it will be high (Feldman & Lewontin, 1975). But, if you compare the skin color of White New Yorkers wintering in Florida to those who winter in New York, there will be a considerable difference in skin color between the groups that has no genetic basis (Feldman & Lewontin, 1975). To claim that racial disparities in education are caused by genetic differences between races on the basis of heritability statistics that are derived within populations, as Jensen (1969) did, is at best a conceptual error (Feldman & Lewontin, 1975), and at worst, an ideological distortion of scientific knowledge (Donovan, 2015a; Graves, 2015; Lewontin, 1996) that is not supported by contemporary population genetic theory on polygenic trait variation (Rosenberg, Edge, Pritchard, & Feldman, 2018).<sup>1</sup>

Altogether, these findings demonstrate that it is incorrect to think that stereotypes are true because individuals of the same race are genetically uniform and people of different races are genetically discrete. Instead, a more biologically accurate view of human difference is that racially defined populations are genetically alike in their variable DNA because each of these populations share similar sets of alleles (Figure 1). That is, we are alike in the ways that we are different.

The purpose of a humane genetics education is to help learners make sense of these complex ideas to reduce their racial biases. Another purpose of a humane genetics education is to help learners understand that claims about race and genetics are not socially neutral. When people claim that economic or educational inequalities between races are the result of genetic differences between races (e.g., Wade, 2014), they often misrepresent, willfully ignore, or misunderstand the social and quantitative complexities of human genetic variation (Coop et al., 2014;

<sup>1</sup>Since the 1930s, population geneticists and anthropologists have argued that cognitive and behavioral traits are malleable because genes produce different trait expressions in different environments due to gene  $\times$  environment interactions (Boas, 1928; Dobzhansky, 1937). Modern research has shown that such traits are shaped through gene  $\times$  environment interactions, epigenetics, environment, and chance (Caspi & Moffitt, 2006; Enoch, 2012; Graves, 2015; Haberstick et al., 2014; Keller, 2014; MacMahon, 1968; Okbay et al., 2016). Moreover, genome-wide association studies have established that only 0.01–0.035% of the variability in educational attainment is associated with any single gene variant (23andMe Research Team et al., 2018; Okbay et al., 2016). The combined contribution of all of these genomic variants to any group level difference in a complex human trait is predicted to be small, possibly spurious, and largely dependent on the environment according to population genetic theory (Rosenberg et al., 2018). Empirically, we also know that intergenerational scores on IQ tests have changed over time because of improvements in the human environment (Bratsberg & Rogeberg, 2018; Flynn, 1999) and the association of IQ with genetic factors depends heavily on environmental factors (Devlin, Daniels, & Roeder, 1997; Flynn, 1999; Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003). To claim that racial inequality in education is immutable because of genetics is no less of a distortion of scientific knowledge today than it was in the 1970s or 1930s.

Donovan, 2015a; Feldman, 2014; Lewontin, 1996). A humane genetics education helps students understand that inequality is not the inevitable product of genes, but it is socially constructed, *partially*, by overly simplified beliefs about genetics (Lewontin, 1996). As prejudice is significantly associated with misunderstandings of human genetic variation (Donovan, 2017; Kimel et al., 2016; Williams & Eberhardt, 2008; Yzerbyt, Corneille, & Estrada, 2001) learning this genetic content for this humane purpose should reduce individual racial bias.<sup>2</sup>

## 2.2 | Reducing racial bias by learning about human genetic variation

The theory of change in a humane genetics education begins with how youth perceive human genetic variation. Studies estimate that 75% of college students taking introductory biology and genetic courses do not know that there is proportionally more genetic variation within ethnic groups than between them (Bowling et al., 2008). Among anthropology students, Hubbard (2017b) found that 29% believed that there is proportionally more biological difference between two races than between individuals within a single race. A study by Donovan (2017) found that students ( $n = 135$ , 7<sup>th</sup>–9<sup>th</sup> graders) attending high socioeconomic status (SES) schools in the San Francisco Bay Area perceived 43% of genetic and phenotypic differences between racial groups and 57% within racial groups. It appears that students perceive far too much genetic variation across races and far too little within races. Due to the large mismatch between students' perceptions of human genetic variation and the actual scientific estimates of such variation, curriculum and instruction that (a) reduces the perception of variation between races and (b) increases the perception of variation within any single racial group should cause reductions in racial bias if such perceptions are causally implicated in the development of racial bias.

The relationship between perceptions of human genetic variation and stereotype endorsement is mediated by genetic essentialism (Dar-Nimrod & Heine, 2011; Donovan, 2015b). Genetic essentialism of race is a social cognitive bias which assumes that the genes *inherent* in people make same race individuals physically and behaviorally *uniform* and people of different races physically and behaviorally *discrete* (Andreychik & Gill, 2014). People who believe in the genetic *uniformity* of same race individuals have been found to believe that stereotypes apply to all group members (Yzerbyt et al., 2001). When people believe that racial groups are biologically *discrete* categories they also tend to endorse racial stereotypes (Bastian & Haslam, 2006) because discreteness beliefs facilitate category-based inductions about group members (Gelman, 2004). Finally, when people believe there are *inherent* differences in the genes of races, they attribute cognitive and behavioral differences between races to genetics (Donovan, 2016, 2017) because believing that groups cohere around inherent characteristics accentuates uniformity and discreteness beliefs (Yzerbyt et al., 2001). Therefore, if learning undermines either the belief that races are genetically discrete or the belief that same race individuals are genetically uniform, it could reduce stereotyping by impacting these essentialist beliefs.

To our knowledge, no randomized trials have tested the hypothesis that constructing an accurate understanding of human genetic variation can cause a reduction in these forms of racial bias. Previous studies have found that when biology education causes adolescents to perceive too much genetic variation between races it also causes an increase in genetic essentialism (Donovan, 2017). Thus, it should be possible to run this process in reverse and reduce racial bias. For instance, Aboud and Fenwick (1999) found that increasing adolescents' perceptions of observable differences (nongenetic) between individuals of the same race can reduce stereotyping among White students who are high in prejudice. Tawa (2016) found that exposing a single cohort of 31 adolescents to a 5-hr intervention that taught about the biological similarity of people was associated with a reduction in genetic essentialism. Yet, as his study was nonexperimental, Tawa (2016) could not establish if his intervention caused these effects. Moreover, his study did not explore whether changes in genetic knowledge

<sup>2</sup>We emphasize that we are not naïve to the fact that the development of racial bias has many causes (and we elaborate on this point in Section 2.3). Rather, we set forth a hypothesis for reducing racial bias in educational settings where such bias is arguably perpetuated through the content of the mainstream biology curriculum (Donovan, 2014, 2015a, 2015b, 2016, 2017).

mediated the declines in genetic essentialism he observed. Hubbard (2017b) found that undergraduates' ( $N = 296$ ) misconceptions about genetic variation and race declined significantly as they learned four key ideas about human genetic difference. But, like Tawa (2016), Hubbard (2017b) could not attribute these declines to her intervention because she used a nonexperimental pre-post design. Nor did she explore whether the declines in misconceptions were associated with changes in students' racial biases.

In sum, extant research and theory are consistent with the claim that genetic essentialism and stereotyping can be reduced by changing perceptions of human genetic variation. How, then, does one design and implement a humane genetics education to accomplish this goal?

## 2.3 | Design and implementation of a humane genetics education

Science educators (e.g., Donovan, 2015b) have situated genetic essentialism within theories of conceptual change (diSessa, 1998; Gregoire, 2003; Heddy & Sinatra, 2013; Pintrich, Marx, & Boyle, 1993; Sinatra, Brem, & Evans, 2008; Sinatra, Kienhues, & Hofer, 2014; Sinatra, Southerland, McConaughy, & Demastes, 2003; Strike & Posner, 1992) because (a) it is a misconception that is negatively associated with biology learning (e.g., see Emmons & Kelemen, 2015; Evans et al., 2010; Opfer, Nehm, & Ha, 2012; Shtulman & Schulz, 2008), and (b) it is a socially motivated form of cognition that people use to make sense of the social world (Dar-Nimrod & Heine, 2011; Keller, 2005) and is therefore motivated by the values and social goals of people (Keller, 2005). However, as critics of conceptual change theory have long pointed out (Pintrich et al., 1993), changing strongly held beliefs—like genetic essentialist beliefs—through education is not a cold and completely rational process and is thus difficult. The difficulty of changing racial beliefs stems from the fact that people develop and use racial beliefs in school.

### 2.3.1 | Schooling and racial beliefs

Racial stereotyping develops, in part, because students are aware that race is salient for making sense of social patterns in school. For instance, studies have found that when students use essentialist beliefs to make sense of racialized patterns observable within school it contributes to stereotype development (Bigler & Liben, 2007; Pauker, Ambady, & Apfelbaum, 2010; Pauker, Xu, Williams, & Biddle, n.d.). This does not mean, however, that racial biases, like stereotypes, are only produced through schooling. Racial bias is a multifactorial trait, which means that multiple causal factors influence its development. Some of the processes influencing racial bias operate from the bottom up in the form of psychological biases such as essentialism (Bigler & Liben, 2007; Cimpian & Salomon, 2014; Pauker et al., 2010). Other factors influence racial bias from the top down. Institutional racism, segregation, the ethnic diversity of one's community, and levels of intergroup contact may interact in complex and unpredictable ways with bottom-up psychological processes to affect people's biases (Carter & Reardon, 2014; Clark, 1991; Emerson, Kimbro, & Yancey, 2002; Pettigrew & Tropp, 2006; Putnam, 2007). However, many developmental studies have found that the interaction of schooling and essentialist thinking is one important process that influences the development of stereotyping. Indeed, developmental intergroup theory (Bigler & Liben, 2007), which is an influential theory of stereotype development because of its broad empirical support, contends that essentialism is an internal cognitive motivation that interacts with racial categorization to affect the development of stereotyping during childhood and adolescence.

Refuting essentialist beliefs about race in school is thus a potentially powerful place to intervene to reduce prejudices in US society as people actively construct explanations for racial difference in school (Bigler & Liben, 2007; Carter, 2012; Donovan, 2014, 2016, 2017; Morning, 2009; Morning, 2011). At the same time, essentialist beliefs in our broader society may be difficult to change via school biology precisely because they are embedded in the context of schools. Indeed, if it is true, as Philip (2011) contends, that racial beliefs are practiced and developed in particular contexts, and are therefore context dependent, then reducing essentialist beliefs within the biology classroom does not guarantee that this reduction will transfer out of the biology classroom and into society at large.

Nevertheless, what we teach students about race in school biology, and how we teach it, does matter. It matters because RCTs have found that when students learn from a standard biology curriculum addressing racial differences in genetic disease prevalence it can *unintentionally* increase students' genetic essentialist biases in ways that appear to affect their attitudes toward racial disparities in education, their explanations of such disparities, and their interest in socializing across racial boundaries (Donovan, 2014, 2016, 2017). And, if biology education is a context where these biases develop unintentionally, then it should also be a place where they are challenged *intentionally*. The real question, then, is how can educators effectively reduce essentialism if the process of reducing it is emotionally "hot" (Pintrich et al., 1993).

### 2.3.2 | A cognitive-affective model for the conceptual change of racial beliefs

Using education to change belief in genetic essentialism is a "hot" process of conceptual change because it makes the learner's racial identity salient to their learning (Gregoire, 2003). In turn, learning experiences that challenge commonsense understandings of racial identity may also challenge other core beliefs that are related to a student's sense of self (Larkin, 2012), thereby creating a perceived ego threat (Cohen & Sherman, 2014) and thus a stressful learning experience for the student. However, under a cognitive-affective model of conceptual change (Gregoire, 2003) the stress induced by a humane genetics education is not necessarily detrimental and it may actually be necessary for a true conceptual change in belief in genetic essentialism of race.

For example, Gregoire's (2003) model predicts that such stress could have two possible impacts on the conceptual change process depending on whether learners respond to a humane genetics education with a challenge appraisal or a threat appraisal. In a challenge appraisal, identity-related stress that is induced by learning will cause students to *systematically process* the information presented to them if students possess the appropriate *motivations* and *prior knowledge* for learning. *Systematic processing* of the information in a humane genetics education will, in turn, increase the probability that students revise their racial beliefs (i.e., accommodation). However, students who have a weak motivation to learn about human genetic variation, and/or those who have insufficient prior knowledge and time to understand it, may perceive identity-related stress as a threat to their ego. In turn, a threat appraisal could lead these students to *heuristically process* the information presented through a humane genetics education. And, if students engage in *heuristic processing*, then there are only two possible outcomes: (a) no belief change or (b) superficial belief change (i.e., assimilation; Gregoire, 2003). Altogether, a cognitive-affective model implies that the stress of a humane genetics education can produce either true conceptual change, superficial change, or no change depending on students' motivations and prior knowledge.

### 2.3.3 | Motivations and prior knowledge

Research suggests that many students likely possess the motivation for true conceptual change of genetic essentialism, but perhaps not the prior knowledge. For example, several studies carried out in the United States and across the world have demonstrated that high school-aged students are most interested, and indeed highly motivated, to learn about human biology (Baram-Tsabari, Sethi, Bry, & Yarden, 2006; Baram-Tsabari & Yarden, 2009, 2010; Gardner & Tamir, 1989; Hong, Shim, & Chang, 1998; Uitto, 2014). So, there is a decent chance that many students will be motivated to learn the content in a humane genetics education. But, unfortunately, studies also show that students possess inaccurate prior knowledge about human genetic variation or no knowledge of it (Bowling et al., 2008; Donovan, 2017; Hubbard, 2017b). This does not mean that students are deficient or to blame, rather it means that biology education currently does very little to teach students about the complexities of human population genetics (Donovan, 2015b), beyond perhaps, Hardy-Weinberg equilibrium or Punnett squares (Dougherty, 2009; Jamieson & Radick, 2013; Stern & Kampourakis, 2017). This oversight is one reason why adolescents are novices in the domain of genetic variation. Consequently, a humane genetics education needs to

account for the fact that high school-aged learners are novices when it comes to learning in the domain of human genetic variation.

One important way to help novices learn about human genetic variation is to select information that fits within their zone of proximal development (Vygotsky, 1978). Indeed, Aboud and Fenwick (1999) argue that educational attempts to reduce racial prejudice in adolescent-aged students are more successful if they include information known to reduce prejudice that fits into the zone of proximal development of students. If too complex, then using scientific information to correct biased thinking can backfire (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). For example, when interventions challenge a myth with a scientific fact, they create an association between the myth and the fact, which increases the risk that students will conflate myth and fact, thereby leading to greater belief in the myth at a later date (Lewandowsky et al., 2012). This backfiring process, which we call the “Lewandowsky effect,” is more probable when interventions seek to replace a simple explanation for the world with a more complex one as we attempt here. The challenge, then, is to select the right information for students to learn, which we have summarized in Figure 1, and then to scaffold that learning to enhance its systematic processing to reduce the probability of intervention backfiring.

### 2.3.4 | Instructional scaffolds implemented to prevent backfiring by enhancing systematic processing

Studies grounded in conceptual change have found that misconceptions can be reduced through a refutational curriculum that triggers a misconception, labels it as incorrect, refutes it with evidence, and provides an alternative way of understanding the phenomena originally explained by the misconception (Guzzetti, Snyder, Glass, & Gamas, 1993; Lewandowsky et al., 2012; Van den Broek, 2010). A well-designed refutational approach does not simply tell learners about how to think about the world differently after a discrepant event; rather it introduces them to a more useful way of understanding the world. For an alternative explanation to be useful, it must be plausible to the learner, which means that it must be congruent with, and connected to, other concepts the learner uses to make sense of the world (Lewandowsky et al., 2012; Vygotsky, 1978). A good alternative explanation also helps the learner understand why prior explanations based in misinformation are flawed and why people tend to believe misinformed explanations (Lewandowsky et al., 2012). When these criteria are met, there is a lower probability of backfiring, which, from the standpoint of conceptual change, means that learners will be less likely to revert to essentialism when making sense of novel race-related social phenomena.

The interventions created for this study were refutational and they included additional scaffolds to enhance systematic processing. For example, after our interventions elicited biological essentialist beliefs about race (i.e., uniformity and discreteness) and labeled them as biologically inaccurate through a discrepant event we then prepared novice learners for an alternative explanation of the discrepant event by using contrasting cases (Schwartz & Bransford, 1998) of genetic data. Contrasting cases were used to scaffold this portion of the intervention because the use of anomalous data to create conceptual change can be ineffective when learners construe data through their preconceptions (Chinn & Malhotra, 2002). After learning from contrasting cases of genetic data, students were told an alternative explanation of racial difference. Then, we used argumentation scaffolds (Osborne, Donovan, Henderson, MacPherson, & Wild, 2016) to support learners in an activity that asked them to critique an essentialist claim about race using their new knowledge of human genetic variation. This step of the learning process further supported conceptual change because critique supports student sense-making in science (Ford, 2012).

The success of this intervention strategy depends, in part, on whether learners will adopt the alternative explanation about racial difference that we offer to them afterward. We offered learners the idea that racial inequality is not a consequence of genetics, rather it is *partially* created through incorrect genetic beliefs. Students were told that Americans develop incorrect ideas about human genetic variation for two *possible* reasons. First, there is a lot of misinformation about race and genetics in our culture (Beckwith, 2013; Lewontin, 1996).

This misinformation has been promulgated by individuals who have a racially biased social agenda (Beckwith, 2013; Lewontin, 1996). Such individuals try to convince other people that racial stereotypes are caused by genetics because they want to mislead people into thinking that racial inequality is natural and therefore not worthy of redress (Graves, 2015; Morning, 2011; Smedley & Smedley, 2005). People who have been misled may then use stereotypes as a way to make sense of the social world around them, further reinforcing the false idea that racial inequality is genetic (Jost, Banaji, & Nosek, 2004; Morin-Chassé, Suhay, & Jayaratne, 2017). Next, we told students that a second *possible* reason why people tend to believe misleading information about the race is because people live in segregated communities in the United States (Allport, 1954; Clark, 1991). As segregation causes social isolation, it prevents people from seeing that individuals in another group are similar to them (Allport, 1954). It could also prevent them from seeing that they differ from people in their own group in the same way that people in another group differ from each other. If individuals never get to have these social interactions, because of segregation, then it will be difficult for them to see that different races are alike because the individuals within races differ in similar ways. Therefore, people will tend to believe misinformation about race and genetics because no one has ever taught them an alternative way of thinking (e.g., Figure 1). This alternative explanation that we offered to students is grounded in social constructionist research about racial bias, such as literature on social-contact theory (Allport, 1954; Pettigrew, 1998; Pettigrew & Tropp, 2006), ideologically motivated reasoning (Morin-Chassé et al., 2017; Morning, 2011), and racial categorization (Chao, Hong, & Chiu, 2013; Williams & Eberhardt, 2008).

### 2.3.5 | The importance of contrasting cases

Adoption of this alternative explanation is more probable if students possess the prior knowledge to systematically process it (Gregoire, 2003) and thus understand it (Alvermann, Smith, & Readence, 1985). Yet learners do not possess well-differentiated prior knowledge about genetic variation that would allow them to understand why essentialism is flawed (Bowling et al., 2008; Donovan, 2017; Hubbard, 2017b) because they are novices in the domain of human genetic variation. We, therefore, designed our intervention using Schwartz and Bransford's (1998) "time for telling" framework to help learners differentiate genetics information that is consistent with, or which conflicts with, genetic essentialism. Schwartz and Bransford (1998) argue that opportunities to analyze information in sets of contrasting cases help learners to perceive features that make a case distinct, thereby helping the novice learner differentiate information like an expert so it can be used in subsequent learning activities, such as reading informational texts (Schwartz & Bransford, 1998) that refute genetic essentialism. We used contrasting cases to help learners understand how genetic data are inconsistent with genetic essentialism so that students could comprehend the complex ideas that were presented in our alternative explanation for racial difference. Table 1 shows an example of this approach.

The contrasting cases introduced learners to counterfactual sets of data that either supported or refuted claims linked to belief in biological essentialism of race. For example, when learners were evaluating a claim about within-group variation, such as "same race individuals are genetically different," they would first be presented with a representation of genetic data that would perfectly support the claim that same race people are genetically different (Table 1A). Then, they would be presented with data in the same graphical format that would perfectly refute this claim (Table 1B). These counterfactuals helped learners to differentiate patterns of human genetic variation by aiding their perception of how data would have to look to be consistent or inconsistent with an essentialist model of human difference. Learners were told that these contrasting cases were fictional and were created to help them make sense of the actual genetic data (Table 1C).

Learners were then given the actual genomic data in the same graphical format along with a text that explained how to interpret it (Table 1C). The real data resembled the antiessentialist counterfactual, but it did not perfectly match it. Instead, the real data shared a deep structure with the antiessentialist counterfactual. Students were thus instructed to argue whether the real data more closely matched the essentialist or

**TABLE 1** Time for telling data interpretation and argumentation scaffold

**A. Counterfactual 1**



**Within group text:** If human genetic data looked like this, then it would *support* the claim that individuals of the same group are genetically different. In Figure 1, we see pie charts for three different ancestral groups. Within each ancestral group about 20% of individuals have each of the five different alleles. Because each ancestral group has five alleles each group has genetic variety.

**Between group text:** Because each group has the same sets of alleles in the same proportions, the groups are genetically identical.

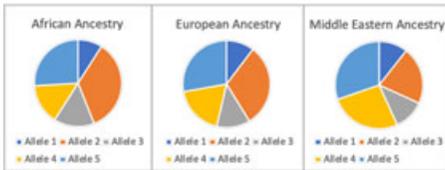
**B. Counterfactual 2**



**Within group text:** If human genetic data looked like this, then it would *refute* the claim that individuals of the same group are genetically different. In Figure 2, we again see pie charts for the same three ancestral groups. In each ancestral group, every individual has the same allele. Since everyone from the same ancestral group would have the same allele, there would be no genetic variety within the ancestral group.

**Between group text:** Since each group has entirely different alleles, the groups are categorically different. They share no genes at this locus.

**C. Argumentation**



**Within group text:** Now let's look at real data from a study on human genetic variation (data from Rosenberg, 2011). Each pie chart shows the different alleles found at the same location on a chromosome within an ancestral group. This pie chart tells us that individuals of African ancestry might have any one of five different alleles at the same location in their DNA (represented by five different colors). The size of each pie slice shows the percent of individuals that have an allele. 33% of individuals with African ancestry have the orange allele, 25% have the light blue allele, and about 5% have the blue allele. The same genetic variability can be seen in people of European ancestry and in people of Middle-Eastern ancestry. **Does the real data support or refute the claim that same race individuals are genetically different?**

**Between group text:** Last time we looked at the number of pie slices in each chart, but this time we want to look at the size of each pie slice. The size of each pie slice is the percentage of individuals in each ancestral group that has a given allele. If you look at all 3 groups, you can see that they all have the dark yellow allele. But, each group differs in the percent of individuals who have the dark yellow allele. About 20% of individuals with African or European ancestry have the dark yellow allele. But, 25% of individuals of Middle-Eastern descent have it. **Does the real data support or refute the claim that different races are genetically alike?**

**D. Time for Telling**



**Within group text - What does this data tell us?** We can see that for each ancestral group, individuals in the group could have any one of several different alleles at a given location on a chromosome. The same pattern is found elsewhere in the human genome. Therefore, scientists have concluded that individuals of the same group are genetically different from each other.

**Between group text - What does this data tell us?** We can see that individuals from different ancestral groups have the same set of alleles at a given location on their chromosomes. But, the percent of individuals in each pie chart who have a specific allele varies from group to group. This means that if you place one pie chart over another they match up very well in color but less well in the size of each pie slice. Thus, different groups are genetically alike because they share similar sets of alleles. In fact, 92.47% of alleles in the variable part of the human genome are shared by two or more groups.

antiessentialist data. Then, learners were asked to use the real data to make an argument either for or against the claim that “individuals of the same race are genetically different” (Table 1C). After, learners were told how to interpret the data (Table 1D).

Using analogical induction to help learners identify the deep structure shared between the antiessentialist counterfactual and real genetic data should result in more efficient learning of the antiessentialist model (see Shemwell, Chase, & Schwartz, 2015). It should also increase the probability that learners transfer this model to other portions of the intervention (see Shemwell et al., 2015). For instance, when exploring between group variation, we asked learners to attend to proportional distinctions across pie charts in Table 1. This provided

learners with an opportunity to see that groups share the same sets of gene variants, but each group is unique because it contains different proportions of people who possess each variant. This sets up a time to tell learners about a new deep structure in the data, which is the idea that different races are genetically alike because they share similar sets of alleles (Figure 1). After helping learners make sense of these ideas, they were told the alternative explanation described above and then asked to engage in the critique portion of the intervention described earlier.

### 2.3.6 | Summary

The instructional framework of a humane genetics education is based on theories of “hot” conceptual change. It embodies the principles of refutational texts, contrasting cases, and scientific argumentation to facilitate students’ systematic processing of human genetic variation data to help students develop the ability to critique genetic essentialist beliefs about race. At present, no studies have demonstrated that learning information about the complexities of human genetic variation will actually reduce racial bias through its impacts on how learners perceive genetic variation. Yet, as we have argued, there are good reasons to think that learning about variation in the domain of genetics can powerfully influence social cognition. At the same time, the cognitive complexities of a humane genetics education lend themselves to misinterpretation in US culture, where ideas about genetic variation and race are usually used to rationalize inequality rather than challenge it (Morning, 2011). Therefore, a humane genetics education could easily backfire (Lewandowsky et al., 2012) and produce more, not less, racial bias. As racial beliefs are constructed and practiced in schools (Bigler & Liben, 2007; Donovan, 2017; Morning, 2011; Willinsky, 1998), it is a worthy cause to explore if racial beliefs can be changed through a more humane genetics education.

## 3 | RESEARCH QUESTIONS AND HYPOTHESES

The main research question driving our study is thus: Does learning scientifically accurate information about genetic variation within and between US census racial groups cause a significant reduction in genetic essentialism and racial stereotyping? Our first two analyses test the predictions that teaching students scientifically accurate information about human genetic variation causes students to: (H1) perceive proportionally less genetic variation between US census races relative to the total variation in humans (i.e., within and between races); and (H2) exhibit less racial bias (i.e., endorsement of genetic essentialism and racial stereotypes). Then, we test the mediational hypothesis that a humane genetics education reduces racial bias by changing perceptions of human genetic variation (H3). Afterward, we explore if reductions in racial bias are reversed over time because of backfiring effects (i.e., Lewandowsky et al., 2012). Specifically, we explore whether students who learn from a humane genetics intervention actually exhibit more racial bias 3 weeks after treatment compared with students who do not receive this intervention.

## 4 | METHODS

To test our hypotheses, we use three different RCTs. We use two different RCT designs, a cross-over trial, and a parallel trial. In RCT 1, we estimate the impact of learning about human genetic variation over an entire week using a group-based version of our intervention through a cross-over trial. Then, in RCTs 2 and 3 we use a 45-min computerized version of our intervention to explore the effects of humane genetics education in a sample of adults and in a sample of adolescents from a geographically diverse set of schools through a parallel trial. Afterward, we meta-analyze the effect sizes from these RCTs to explore their reproducibility. Then, we test whether a single mediating mechanism can account for the findings in each RCT. Finally, we test for “Lewandowsky effects” by exploring the duration of treatment effects over 3 weeks using a subset of our third sample ( $n = 266$ ), under the

assumption that the presence of a Lewandowsky effect would be associated with a reversal in treatment effects several weeks following the intervention.

## 4.1 | Sample descriptions of RCTs 1–3

### 4.1.1 | RCT 1

In the spring of 2016, we recruited students ( $n = 166$ ) from a high SES private middle school ( $n = 52$ , 8<sup>th</sup> graders) and a high SES public high school ( $n = 114$ , 9<sup>th</sup> graders) in the San Francisco Bay Area ( $M_{\text{age}} = 14.3$ ,  $SD_{\text{age}} = 0.74$ , % female = 53%). A minority of students (6.8%) at the high school were eligible for free and reduced-price lunch (FRPL). The sample self-identified as White (48%), Mixed-race (19.88%), Asian (18.1%), and Hispanic (5.42%), with fewer students identifying as Black and Pacific-Islander (1.2% each). The remaining students declined to pick one of the US census categories (6.02%). Students were not paid for their efforts.

### 4.1.2 | RCT 2

The purpose of RCT 2 was to pilot the computerized version of our intervention to test if it worked as intended before using it with students. In the winter of 2018, we recruited adults ( $n = 176$ ) from Amazon's Mechanical Turk website. Participants were paid \$9.00 for their efforts. The sample of adults self-identified as White (76.7%), Black (13.6%), Mixed-Race (3.9%), Hispanic (3.9%), and Asian (1.7%). The majority of participants self-identified as female (63%) and the average age was 41 years old ( $SD = 12.24$ , range 21–73). Participants identified their political orientation as Republican (26%), Independent (24.43%), and Democrat (49.57%). Participants had either never finished high school, graduated high school but did not finish college, or earned associate degrees.

### 4.1.3 | RCT 3

This sample of 9<sup>th</sup>–12<sup>th</sup> graders ( $n = 721$ ) was recruited from five schools in the spring of 2018. Two were public high schools in major cities in Colorado (33.34%). In California, the one public high school was located in the San Francisco Bay Area (41.19% of the sample; same school in RCT 1 but different students). We also sampled one public high school in the Greater Boston, Massachusetts area (20.39% of the sample) and one private high school in the Washington, DC area (4.99% of the sample). FRPL in the two Colorado high schools was high at one site (FRPL = 66%) and low at the other (FRPL = 12.1%). FRPL for the remaining schools was low (Washington DC = 0%, Boston = 25.9%, and California = 6.8%). The percent of White students at each school ranged from 71% in Boston to 20% at one of the Colorado sites. Students self-identified their race as White (61.7%), Asian (19.8%), Mixed-Race (9.9%), Hispanic (4.9%), Black (2.4%), Pacific-Islander (0.55%), and American Indian (0.4%). The mean age of students was 15 ( $SD = 1.02$ , range 14–18) and 47.7% identified as female. Students were in 9<sup>th</sup> grade (54.1%), 10<sup>th</sup> grade (19.6%), 11<sup>th</sup> grade (25.9%), and 12<sup>th</sup> grade (0.4%). We explore backfiring using data from our California school site because this was the only district that agreed to allow us to collect a delayed posttest.

### 4.1.4 | Sample summary

We attempted to recruit twice as many schools, especially those serving lower income students, more racially diverse populations, or those located in more politically conservative areas. However, district research offices for these schools rejected our study. One district research coordinator even rejected the study in an email because its intended purpose was to assess “psychological bias based on racial stereotyping.” Although this sample is geographically diverse, it is predominantly high SES, majority White, and includes schools that had district research priorities that were aligned with our research goals. The particularities of our samples are nevertheless important from a theoretical perspective. Stereotypes about the intellectual abilities of non-White individuals are often used

to deny people social and economic opportunities in high SES communities (e.g., Fryer & Levitt, 2004; Morton, Hornsey, & Postmes, 2009). Moreover, there are implicitly racist beliefs in the majority White science teaching population and teachers' beliefs about the abilities of racially "othered" students negatively impact these students (Bryan & Atwater, 2002). For instance, African Americans earn fewer graduate and undergraduate science, technology, engineering and mathematics (STEM) degrees in fields where more professors, postdocs, and students believe that science ability is inherited (Leslie, Cimpian, Meyer, & Freeland, 2015; Storage, Horne, Cimpian, & Leslie, 2016). As African Americans are stereotyped as lacking a genetic potential for intelligence because people harbor essentialist biases about race, genetic essentialist thinking among teachers and students in the STEM pipeline may be one reason, *among many other reasons*, for African American attrition in STEM education. Altogether, it is important to explore if biology education can reduce racial bias in racially and socioeconomically privileged populations because scientists and science teachers are traditionally from racially and economically privileged groups (Aikenhead, 1996; Bryan & Atwater, 2002; Taie & Goldring, 2018).

## 4.2 | Treatments used in RCTs 1–3

### 4.2.1 | RCT 1

A full description of the race and human genetic variation intervention used in RCT 1 can be found in the Supporting Information Materials. In brief, the RCT 1 intervention taught students about the amount of genetic and phenotypic variability within US census races and between them. It also taught students how people tend to misrepresent those differences when claiming that race is biologically real. In contrast, the climate variation control condition taught students about the amount of temperature and precipitation variability within and between climate zones. Likewise, it taught students how people tend to misrepresent those differences when claiming that climate change is not real. Each intervention was designed to align with the core ideas, cross-cutting concepts, and practices of the Next Generation Science Standards (2016).

Table 2 provides an overview of how the treatments differed in RCT 1. Both treatments used the same instructional framework (outlined in our conceptual framework) and both involved the same time spent on each task. The two interventions differed only in content objectives. We used a unit on climate variation for a control condition for two reasons. First, this unit taught about the concept of variation just like the race condition. Second, it is a politically controversial issue just like the topic of racial difference. This treatment-control contrast ensures that both groups learn about controversial topics in the media associated with scientific estimates of variation, thereby controlling for cognitive and political confounding. This is important because biological beliefs about race are ideological issues (Lewontin, 1996). To not control for ideology would undercut our claim to have run RCTs.

It might be argued that a more clinically relevant control condition would involve business-as-usual (BAU) materials involving race and genetics, such as learning about the prevalence of genetic diseases in various racial groups. However, previous studies have found that such materials increase genetic essentialism (Donovan, 2014, 2016, 2017). Other studies have found that merely exposing individuals to genetic information primes belief in genetic essentialism (Lynch, Bevan, Achter, Harris, & Condit, 2008). As we wanted to have an inert control, we did not use either of these BAU controls. Arguably, the effects we estimate in this paper are smaller than those that will be found if our intervention is compared with such BAU controls in the future.

### 4.2.2 | RCTs 2–3

For these studies, we distilled the ideas in Table 2 into two 45-min computer-based interventions using the learning approaches laid out in our conceptual framework (Table 1). The interventions were delivered through the Qualtrics platform. These computer-based interventions were vetted by the advisory board of our NSF grant, which included a population geneticist, sociologist, educational methodologist, developmental psychologist, two science education professors with expertise in functional and sociolinguistics, and a

**TABLE 2** Conceptual differences between treatments

	Core Ideas Taught	Day	Core Ideas	NGSS standards
<i>Human Variation</i>	1. Scientists do not agree about whether race is biologically real.	1.	1	LS3A
	2. 99.9% of the DNA between any two humans is identical. When geneticists look at the variable portion of human DNA (0.1%) they find:	2.	2.i.	LS3B
	(i) 95.7% of differences are between people of the same race	3.	2.ii., 3	LS4A
	(ii) 4.3% of differences are between people of different races	4.	2.i, 2.ii	
	3. Skin color changes continuously as one moves away from the equator. But there is more variation in skin color across races than within races.	5.	4	
	4. When people make arguments about the superiority of one race over another, they tend to overestimate the amount of genetic difference between races.			
<i>Climate Variation</i>	1. Scientists agree that the climate is changing.	1.	1, 2	ESS2D
	2. Weather and the climate are different concepts.	2.	2, 3	ESS3D
	3. When scientists support claims about climate change they use data on climate variation and not weather variation. For example, if we look over the last 100 years:	3.	2, 3	
	(i) We can see that daily and monthly temperatures and precipitation change – this weather variation occurs within a climate zone and cannot be used to evaluate claims about climate change	4.	4	
	(ii) But, when we look across large land areas and periods of time greater than 30 years we see a continuous increase in the average temperature and precipitation across the United States – this climate variation data can be used to evaluate claims about climate change	5.	4	
	4. When people make arguments that climate change is not real, they tend to incorrectly use evidence about the weather to evaluate claims about climate.			

Note. NGSS: Next Generation Science Standards.

professor of gender and equity studies. Our materials were also vetted by three cooperating teachers and we piloted them with 8<sup>th</sup> graders and 12<sup>th</sup> graders by performing think-alouds. This process resulted in two interventions targeting the core ideas in Table 2 that were written at the same grade level (8<sup>th</sup> grade). Links to the actual computerized interventions can be found in the Supporting Information. To summarize, the computerized interventions informed students about all of the information in Table 2 using the learning approaches in Table 1, just like the group-based interventions in RCT 1.

### 4.3 | Dependent variables in RCTs 1–3

#### 4.3.1 | Timing of measurements in RCTs 1–3

In RCT 1 all instruments were administered at 1-week intervals through a computer-based survey. We measured the students at baseline in RCT 1 to estimate the psychometric properties of our instruments and establish their convergent and discriminate validity. In RCTs 2–3, we only measured the two dependent variables after individuals were treated with the interventions to avoid Solomon effects (Solomon & Lessac, 1968). In RCT 3 at our California school site, we were able to collect an additional delayed postmeasurement of perceptions of human genetic

variation and racial bias 3 weeks after students were treated to explore intervention backfiring via Lewandowsky effects.

### 4.3.2 | Perceptions of human genetic variation (H1)

We used the perceptions of biological variation measure (RCT 1:  $\alpha = 0.90$ ; RCT 2:  $\alpha = 0.88$ ; RCT 3:  $\alpha = 0.86$ ) which was developed and validated to measure adolescent perceptions of genetic variation within and between races (Donovan, 2017). In the first four items of this measure, students are presented with a  $10 \times 10$  matrix of purple circles that represent 100 people of the same race (labeled as either "White," "Black," "Asian," or "American Indian"). The dots in these diagrams changed color randomly to become more heterogeneously colored as students moved a slider bar under the picture on a scale of 0–100% (more color variation = more within-group variation = higher score). Students were asked to move the bar to a location to represent the percentage of DNA that differed between individuals of the same race. Next, students were presented with a series of six overlapping Venn diagrams, one set for each combination of the four US Census races listed above, which could be moved over each other to represent the biological difference between two groups. Students were instructed to make the Venn diagram represent the percentage of genetic material shared between the two labeled groups. Less overlap of the circles equated with a higher percent difference.

A principal components analysis supplemented with a parallel analysis demonstrated that these two item types, between and within items, were statistically discriminable (see Supporting Information). As in previous studies (Donovan, 2017), we averaged the between race and within race questions separately. Then, we applied Equation (1) to these data, which yielded a single proportion for each student that could take on any value between 0 and 1.

$$\frac{\text{Perceived differences between races}}{\text{Perceived differences between races} + \text{Perceived differences within races}} \quad (1)$$

Equation (1) implies that students can increase in score on the instrument either by making the dot diagrams less colorful and/or by placing the circles in the Venn diagrams further apart. Higher scores on this instrument (0–100%) indicate that a student perceives a greater proportion of differences between races relative to the total variation perceived within and between races. Before the start of the experiment, students in RCT 1 perceived, on average, 42% of the human genetic variation between races ( $SD = 24\%$ ). Adults in the control condition of RCT 2 perceived 37% of the human genetic variation between races ( $SD = 27\%$ ). Students in the control condition in RCT 3 perceived 36.5% of the human genetic variation between races ( $SD = 28\%$ ).

### 4.3.3 | Racial bias (H2)

To reduce Type I error we created a composite measure that combined a racial stereotyping instrument (see Levy, Stroessner, & Dweck, 1998) with items from two different genetic essentialism of race instruments (see Parrott et al., 2005; Williams & Eberhardt, 2008) to assess racial bias (as defined in Engberg, 2004). The instrument (RCT 1:  $\alpha = 0.85$ ; RCT 2:  $\alpha = 0.85$ ; RCT 3:  $\alpha = 0.84$ ) assessed agreement with items such as: "Racial groups are primarily determined by biology," "two Black people will always look more similar to each other than a Black person and a White person ever would," "Racial differences in academic ability are caused by genetics," "A randomly picked black person will be unintelligent." Items in each scale were standardized (Z-scored) and then averaged into a single variable of racial bias. Higher scores indicate greater bias. For example, in the Supporting information, we show that students in RCT 1 scoring higher on this measure exhibited significantly greater levels of ingroup favoritism and intergroup anxiety. They were also more certain that they knew people who believed that racial inequality is natural, cannot change, and is not worthy of governmental redress. As expected, students scoring higher on this

measure also perceived significantly more genetic variation between racial groups in RCT 1 ( $r = 0.19$ ,  $p = 0.017$ ), RCT 2 ( $r = 0.26$ ,  $p = 0.016$ ), and RCT 3 ( $r = 0.13$ ,  $p = 0.017$ ).

#### 4.4 | Cross-over trial design of RCT 1

Individual students were randomly assigned to learn about human genetic variation (the treatment) or climate variation (the control) for five, 45-min, periods. We randomly assigned students within classrooms into two new classrooms so that students could learn in groups without introducing cluster level-correlations that reduce statistical power. Table 3 describes the design.

Formally, this design is called an individually randomized trial with clustering (IRTC) because individuals are randomized to experimental arms and then treated as a group (Kahan & Morris, 2013; Weiss, Lockwood, & McCaffrey, 2016). We conducted a cross-over IRTC where half of the students were assigned to receive the human genetic variation and race intervention first, then the climate variation intervention second (R1C2). The other half of the students received the climate variation intervention first, then the race intervention second (C1R2).

After being randomized into treatment groups, students learned in separate classrooms for 2 weeks at each school. Random assignment created a total of 16 new classrooms. Because two different teachers implemented both interventions at each time point, we include fixed effects for teacher in our statistical modeling to control for time-invariant differences between teachers.

#### 4.5 | Parallel trial design of RCTs 2 and 3

In our two computer-based RCTs, we block randomized individuals into experimental arms within self-identified racial groups. In RCT 3 there was an additional layer of block randomization because students were randomized to experimental arms within classrooms by self-identified race. Therefore, RCTs 2 and 3 differ from RCT 1 because there is no cross-over trial. Participants in these two trials were randomly assigned to learn either about race or climate and then they completed the surveys of the dependent variables. They did not receive the alternate intervention at a later date. Table 3 summarizes the design of this parallel trial.

#### 4.6 | Analysis summary

We report treatment effects on Z-scores of all variables, except in the mediation analyses. Standard errors are calculated at the person level because each RCT was person-randomized and there was a low intraclass correlation (ICC) across classrooms for both dependent variables (e.g., before treatment in RCT 1,  $ICC = 0.02$ – $0.071$  and in RCT 3,  $ICC = 0.009$ – $0.066$ ). In the Supporting Information, we further describe our statistical methods and demonstrate that random assignment in each RCT produced baseline equivalence. There we also show that there was no attrition and that missing data were missing at random. We also report findings from models using multiple imputation.

In RCT 1, where we test H1 and H2, we report intention to treat estimates from marginal regressions that modeled within-subject correlations through an unstructured covariance matrix. To test the reproducibility of the findings supporting H1 and H2 observed in RCT 1 we compare them with the findings of RCTs 2 and 3 using a DerSimonian and Laird (1986) random effects meta-analysis. We took a random effects approach to meta-analysis because our three RCTs involved different versions of the intervention and they were conducted in different populations. The details of our meta-analysis (including all statistical code) are described in Kontopantelis and Reeves (2016). For our mediation hypothesis (H3) we use a Sobel–Goodman test augmented with bootstrapping ( $n = 5,000$  replications) in each RCT. Finally, to explore backfiring we test if effects reverse after 3 weeks using a subset of our third sample and multivariate regression.

**TABLE 3** Experimental designs of RCT 1–3

A. Cross-over design of RCT 1							
School	Randomly Assigned Groups	Day 0	Intervention at Time 1 (T1)	Day 7	Intervention at Time 2 (T2)	Day 14	Day 21
Middle School	Race at T1 & Climate at T2 (R1C2)	Measure 1	Race (Teacher1)	Measure 2	Climate (Teacher1)	Measure 3	Measure 4
	Climate at T1 & Race at T2 (C1R2)		Climate (Teacher2)		Race (Teacher2)		
High School	Race at T1 & Climate at T2 (R1C2)		Race (Teacher3)		Climate (Teacher3)		
	Climate at T1 & Race at T2 (C1R2)		Climate (Teacher1)		Race (Teacher1)		
B. Parallel Design of RCTs 2-3							
Occurred within a single one-hour block period							
Within each classroom	Students randomly assigned to one of two interventions	Race Intervention			Post-Test of Dependent Variables		
		Climate Intervention					

Note. RCT: randomized controlled trial.

## 5 | RESULTS

### 5.1 | Initial support for H1 and H2 in RCT 1

We predicted that the race intervention would cause declines in perceptions of human genetic variation and racial bias. Compared with the climate variation intervention (C1R2), the race variation intervention (R1C2) caused greater declines in the perception of genetic variation between races,  $\chi^2(1) = 7.10$ ,  $p = 0.007$ , and in levels of racial bias,  $\chi^2(1) = 9.57$ ,  $p = 0.002$ ;  $ps < 0.05$  (after Bonferroni adjustment between time points 1 and 2; Figure 2a,b). Thus, as predicted, we found significantly greater declines in perceptions of genetic variation and in racial bias when comparing students in the race intervention to students in the climate intervention. After treatment at time point 2, students in the race condition (R1C2) also perceived significantly less genetic variation between races ( $\beta = -0.524$ , standard error [SE] = 0.191,  $p = 0.006$ ,  $d = -0.59$ ) and also exhibited significantly less racial bias ( $\beta = -0.308$ ,  $SE = 0.128$ ,  $p = 0.016$ ,  $d = -0.48$ ;  $ps < 0.05$ , after Bonferroni adjustment) than students in the climate condition (C1R2; Figure 2a,b, time point 2).

When the second half of students received the race variation intervention, they also declined significantly more in the perception of genetic difference,  $\chi^2(1) = 21.7$ ,  $p = 0.0001$ , and in racial bias ( $\chi^2(1) = 31.67$ ,  $p = 0.0001$ ) ( $ps < 0.05$ , after Bonferroni adjustment), compared with students who received the climate variation intervention second (compare C1R2 with R1C2 between time points 2 and 3 in Figure 2a,b). Thus, our results are replicated in our cross-over design of RCT 1. Teaching these students about human genetic variation caused reductions in racial bias (H2) and the perception that races are genetically different (H1).

### 5.2 | Reproducibility of H1 and H2 findings in RCTs 2 and 3

Having established support for our hypotheses in RCT 1, we then explored the reproducibility of the study by conducting RCTs 2 and 3. We then performed a statistical meta-analysis to combine the estimates of our three studies, leveraging and preserving the randomization characteristics of each study in the combined estimate. Figure 3 represents the results of a DerSimonian and Laird (1986) random effects meta-analysis of the treatment effects of the race intervention on the variables used in each of the three RCTs.

### 5.2.1 | Perceptions of human genetic variation (H1)

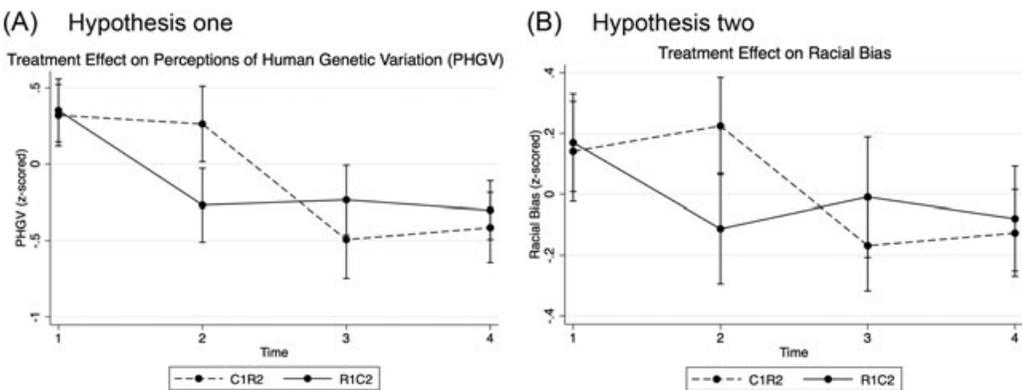
As Figure 3 shows, RCTs 2 and 3 created similar magnitude reductions in perceptions of human genetic variation as those observed in RCT 1 (Cochrane's  $Q = 1.72$ ,  $df = 2$ ,  $p = 0.424$ ,  $I^2 = 0\%$ ). On average, across all three RCTs, learning about human genetic variation results in a three-quarters standard deviation reduction in perceptions of human genetic variation (Cohen's  $d = -0.755$ , 95% confidence interval [CI]  $[-0.876, -0.634]$ ).

### 5.2.2 | Racial bias (H2)

Figure 3 also shows that across all three RCTs, learning about human genetic variation through our interventions caused a significant mean reduction in racial bias (Cohen's  $d = -0.425$ , 95% CI  $[-0.721, -0.129]$ ). But there was significant variability across studies in the magnitude of this effect. Both the  $Q$ -test and the value of  $I^2$  indicate true between-study variation in effect sizes (Cochrane's  $Q = 8.64$ ,  $df = 2$ ,  $p = 0.013$ ,  $I^2 = 76.84\%$ ). That is, the populations in the three studies do not share a common effect size. The reductions in racial bias appear greater in RCT 2 ( $d = -0.663$ , 95% CI  $[-0.957, -0.369]$ ) and RCT 1 (Cohen's  $d = -0.476$ , 95% CI  $[-0.786, -0.166]$ ) compared with RCT 3 ( $d = -0.203$ , 95% CI  $[-0.350, -0.056]$ ). Put differently, the race intervention caused a statistically significant reduction in racial bias in each study (relative to the climate intervention), but the magnitude of the reduction varied across RCTs.

### 5.3 | Support for H3 in RCTs 1–3

Having found evidence in support of our first two hypotheses, we then explored whether the effect of the race intervention on racial bias was transmitted by changes in perceptions of human genetic variation (H3). Figure 4 summarizes the results of these tests. In RCT 1, we find that 22.5% of the total reduction in racial bias caused by the race intervention is transmitted through perceptions of human genetic variation (bootstrapped indirect effect:  $\beta = -0.093$ , 95% CI  $[-0.201, -0.028]$ ). In RCT 2, we find that 18.6% of the total reduction in racial bias caused by the race intervention is transmitted through perceptions of human genetic variation (bootstrapped indirect effect:  $\beta = -0.053$ , 95% CI  $[-0.129, -0.011]$ ). In RCT 3, we find that 44.8% of the total reduction in racial bias caused by the



**FIGURE 2** Treatment effects for confirmatory analyses. C1R2 and R1C2 refer to the timing of the climate and race interventions for each group, respectively. An understanding of that timing is essential to the interpretation of these graphs. Time point 1 marks the pretest for each group. Between time points 1 and 2, the R1C2 group received the race intervention and the C1R2 group received the climate intervention. Between time points 2 and 3 the R1C2 group received the climate intervention and the C1R2 group received the race intervention. Time point 4 marks the delayed posttest for each group [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

race intervention is transmitted through perceptions of human genetic variation (bootstrapped indirect effect:  $\beta = -0.086$ , 95% CI [-0.148, -0.032]). As these effects are not significantly different, learning about genetic variation appears to reduce racial bias by changing how adolescents perceive human genetic variation.

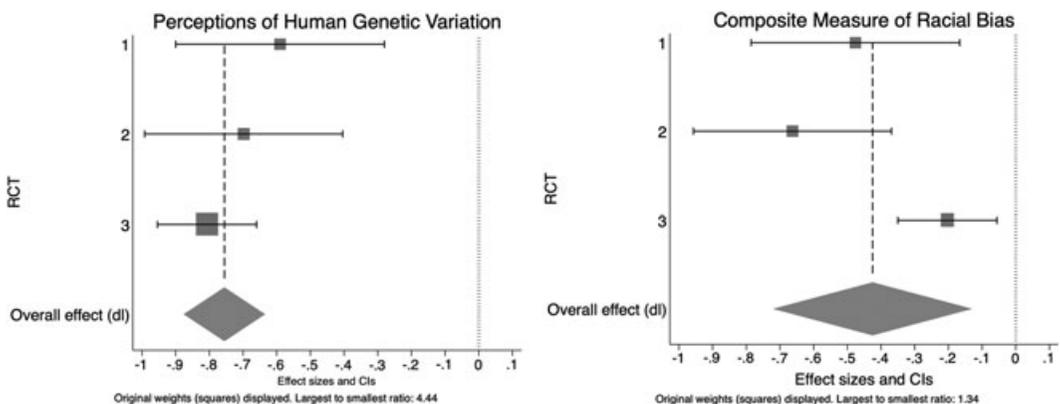
To understand how this occurred, we reran each mediation model using a disaggregated form of the perceptions of human genetic variation instrument, using only the between-group items or only the within-group items. In RCTs 1–3 we found no evidence of an indirect effect using the within-group items ( $-0.030 \leq \beta \leq -0.00049$ ,  $ps \geq 0.28$ ) but we did find evidence of an indirect effect using the between-group items ( $-0.126 \leq \beta \leq -0.052$ ,  $ps \leq 0.019$ ). Therefore, learning about human genetic variation appears to reduce racial bias by affecting how students perceive the discreteness of racial categories rather than how students perceive the uniformity of same race individuals.

#### 5.4 | Duration of effects in RCT 3

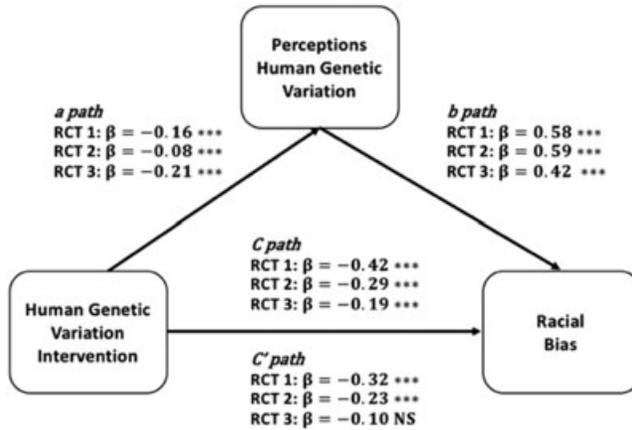
Having found evidence in support of the humane genetics education hypothesis we then explored whether the effects on racial bias persisted over time. We collected a delayed postmeasurement of both dependent variables 3 weeks after treatment at our California school site (RCT 3,  $n = 266$ ). Using ordinary least squares regression, we found that students in the race condition still perceived significantly less variability between US census races than students assigned to the climate condition ( $\beta = -0.09$ ,  $SE = 0.032$ ,  $p = 0.008$ ) and also had significantly lower levels of racial bias ( $\beta = -0.26$ ,  $SE = 0.095$ ,  $p = 0.006$ ). We then conducted a time-lagged mediation model to explore if changes in perceptions of genetic variation measured immediately after treatment mediated reductions in racial bias 3 weeks later. We found that 32% of the treatment effect on racial bias in the delayed post was mediated by changes in how students perceived genetic variation immediately after they were treated 3 weeks earlier (indirect effect:  $\beta = -0.078$ ,  $SE = 0.039$ ,  $p = 0.048$ ). As these effects were not reversed, there is no evidence that the race intervention backfired due to a “Lewandowsky effect.” In the Supporting Information, we support this claim further by using a multilevel Bayesian analysis of our data.

#### 5.5 | Summary of results

Table 4 maps all findings from all three RCTs to our hypotheses and research questions. The summary of results in Table 4 clearly demonstrates that our interventions reduced racial bias by changing how learners perceived human



**FIGURE 3** Forest plot of random effects meta-analyses of RCTs 1–3. The gray diamond refers to the meta-analytic mean effect and confidence interval of the race intervention on each variable. Gray boxes and confidence intervals show the mean effect of the intervention in each RCT. Confidence intervals that do not overlap zero show a significant effect at  $p < 0.05$ . For RCT 1, we use the marginal treatment effects from time point 2 in Figure 2. RCT: randomized controlled trial [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 4** Exploratory tests of the mediated treatment effect (Hypothesis 3). ns: not statistically significant. \*\*\* $p < 0.001$ . For RCT 1 we tested for mediation using the data only from the first half of the cross-over trial at time point 2. RCT: randomized controlled trial

genetic variation and these effects lasted for at least 3 weeks in our samples of predominantly White, high SES students.

## 6 | DISCUSSION

Across three RCTs we demonstrate that when students (RCTs 1 and 3) and adults (RCT 2) were taught scientifically accurate information about human genetic variation, it reduced their racial bias and the perception that races are genetically different. These findings establish proof of concept for the hypothesis that human genetics education can be used to create a more humane society (i.e., less prejudiced) by influencing social cognition. We predicted these effects on the basis of prior research which suggested that learning about human genetic variation could reduce racial bias either by undermining the belief that same race people are genetically uniform and/or the idea that racial groups are discrete and nonoverlapping categories. Our mediation analyses suggest that racial bias was reduced by our interventions primarily because individuals perceived less between group variation rather than more within-group variation. However, this finding does not mean that learning about within-group variation is unimportant as the two forms of variation are proportionally related. A learner must make sense of within-group variation if they are to understand between group variation or the idea that racial groups are genetically alike in their variable DNA because they share similar sets of within-group differences (Figure 1). Constructing an understanding of these concepts together reduced racial bias.

To elaborate further, when people are led to believe that scientists can distinguish one group from another based on traits that are similar within a group and different between groups, they are more likely to search for an underlying essence to differentiate ingroups from outgroups (Yzerbyt et al., 2001). Increased belief in an inherent essence then leads individuals to evaluate ingroups more favorably than outgroups (Yzerbyt et al., 2001) and to categorize people into more discrete racial groups (Chao et al., 2013). In turn, a belief in racial discreteness has been found to lead people to stereotype outgroups more strongly because it causes them to perceive illusory correlations between outgroups and traits (Yzerbyt et al., 2001). Arguably, we ran this mechanism in reverse. By helping learners construct an understanding of genetic variation within and between groups, our interventions changed how students perceived between group variation, which reduced racial bias by undermining belief in racial discreteness.

An avenue for future research that could further explore the plausibility of this explanation could be to investigate the role that visualization and representation played in these results. Our studies used an instrument that required students to visually represent their mental models of human genetic variation using dot diagrams (for within-group variation) and Venn diagrams (for between-group variation). Furthermore, our race interventions used these same representations to help students learn about genetic variation. Although a visual approach to representing genetic variation is common in population genetics research (Donovan, 2017), the representations of variation that we used in our study do not appear to be a common component of teaching about genetic variation. Instead, students tend to learn models of genetic variation with numbers (e.g. Hardy-Weinberg equilibrium) or matrices (e.g., Punnett squares) during BAU biology instruction (Stewart, Cartier, & Passmore, 2004). However, studies have found that instruction requiring students to engage in visualization through representation can enhance students' model-based reasoning and biology learning (Ainsworth, Prain, & Tytler, 2011; Dauer, Momsen, Speth, Makohon-Moore, & Long, 2013; Lehrer & Schauble, 2004; Ryoo & Linn, 2012; Wilkerson & Laina, 2017; Wilkerson-Jerde, Gravel, & Macrander, 2015). Thus, the effects we observed in these studies may have only occurred because students visually represented their mental models of genetic variation. A  $2 \times 2$  factorial study of our race intervention that manipulates how students represent their understanding of variation (Factor 1: with numbers or with drawings) and how they learn about variation (Factor 2: with numbers or with drawings) could shed light on the extent to which our results depended upon students visually modeling genetic variation.

Another cognitive ambiguity in these findings that deserves further study is whether a humane genetics education produces true conceptual change (i.e., accommodation) or just superficial conceptual change (i.e., assimilation). True change would manifest across social contexts. It would also require an empirically grounded justification using a specific model of conceptual change, such as the cognitive-affective model we used here (Gregoire, 2003). This justification would necessitate that future studies of a humane genetics education measure the stress level of students, their perception of threat, and their prior knowledge and motivations to see how each interacts with the content in a humane genetics education to affect students' racial beliefs. We did not measure these constructs and we did not show that our interventions produced impacts on belief in genetic essentialism observable across social contexts. Therefore, while we can claim that we significantly reduced belief in genetic essentialism we cannot be sure how much of this reduction was the product of true conceptual change versus superficial change.

Another ambiguity in our findings is how instructional delivery and sampling characteristics interact to affect their replicability. We used two different kinds of interventions that differed because they involved or did not involve group-based discussions. We then tested these two different instructional approaches on two different populations (adolescent and adult). The invariance of the treatment effects on perceptions of human genetic variation across all three RCTs suggest that a humane genetics education can reduce the perception of genetic variation between US census races in samples of similar demographics drawn from culturally similar schools. These changes in perceptions of human genetic variation should be possible if teachers use curriculum and instruction that is aligned with the learning theory and ideas of a humane genetics education.

Whether teachers can reproduce the effect on racial bias is somewhat less certain than whether they can reproduce the effect on perceptions of human genetic variation. Our longer group-based intervention and shorter computer-based intervention both caused significant reductions in racial bias among adolescent learners (RCTs 1 and 3) but there was significant treatment effect variability in the magnitude of these reductions. For instance, the reductions in RCT 1 were greater than the reductions in RCT 3. Aboud and Fenwick (1999) argue that school-based debiasing interventions work better when high and low prejudice students talk together. They found that when high prejudice students made stereotypical claims about a group, low prejudice students would propose counterexamples to the stereotype during classroom discussions. In this zone of proximal development, high prejudice students revised their thinking through social interaction. Students in RCT 1 argued about racial difference using genetic evidence in small groups, whereas students in RCTs 2 and 3 did not. Therefore, the cognitive work of learners during a humane genetics education arguably needs to be mediated by teachers and

**TABLE 4** Summary of results from all randomized controlled trials (RCTs)

	RCT 1	RCT 2	RCT 3	Summary
Sample Description	52 eighth grade students & 144 ninth grade students learning together in classrooms for 5 days.	176 adults learning online for 45 minutes by themselves.	721 high school students learning online for 45 minutes by themselves within a classroom.	Predominantly white, high SES samples.
H1: The race intervention will cause learners to perceive less genetic variation between U.S. census races.	In between time points 1-2, there was a significantly greater average decline in the perception of genetic variation between races in students who learned from the race intervention compared to those who learned from the climate intervention ( $\chi^2(1) = 7.10, p = 0.007$ ).	After treatment, adults' perceptions of between group genetic variation were significantly lower in the race condition compared to the climate condition, on average ( $\beta = -0.08, p < 0.001$ ).	After treatment, students' perceptions of between group genetic variation were significantly lower in the race condition compared to the climate condition, on average ( $\beta = -0.21, p < 0.001$ ).	The race intervention caused lower perceptions of human genetic variation in all three studies.
H2: The race intervention will cause a reduction in racial bias in those who learn from it.	In between time points 1-2, there was a significantly greater average decline in racial bias in students who learned from the race intervention compared to those who learned from the climate intervention ( $\chi^2(1) = 9.57, p = 0.002$ ).	After treatment, adults' levels of racial bias were significantly lower in the race condition compared to the climate condition, on average ( $\beta = -0.29, p < 0.001$ ).	After treatment, students' levels of racial bias were significantly lower in the race condition compared to the climate condition, on average ( $\beta = -0.19, p < 0.001$ ).	The race intervention caused lower racial bias in all three studies.
H3: The race intervention will reduce racial bias by changing how learners perceive human genetic variation.	22.5% of the reduction in racial bias caused by the race intervention was transmitted by reductions in perceptions of human genetic also caused by the intervention (indirect effect: $\beta = -.093, 95\% \text{ CI } [-.201, -.028]$ ).	18.6% of the average difference in racial bias caused by experimental conditions was transmitted by intervention induced differences in perceptions of human genetic variation (indirect effect: $\beta = -.053, 95\% \text{ CI } [-.129, -.011]$ ).	44.8% of the average difference in racial bias caused by experimental conditions was transmitted by intervention induced differences in perceptions of human genetic variation (indirect effect: $\beta = -.086, 95\% \text{ CI } [-.148, -.032]$ ).	In all studies, the race intervention decreased racial bias by changing how learners perceived human genetic variation.
Backfiring: Do the above effects reverse over time?	The reductions in perceptions of human genetic variation and racial bias persisted one week after both groups of students were treated (see Figure 2, time point 4).	Not Applicable	For at least 3-weeks after intervention the average differences in racial bias ( $\beta = -.26, p = 0.006$ ) and perceptions of genetic variation ( $\beta = -.09, p = 0.008$ ) were maintained. Also, the mediation effects persisted for at least 3-weeks after intervention ( $\beta = -.078, p = 0.048$ ).	There is no evidence in these two studies to suggest that they do.
Reproducibility: Does the race intervention cause similar sized reductions in perceptions of human genetic variation and racial bias relative to the climate intervention?	The effect size for perceptions of human genetic variation ( $d = -.590$ ). The effect size for racial bias ( $d = -0.476$ ).	The effect size for perceptions of human genetic variation ( $d = -.698$ ). The effect size for racial bias ( $d = -0.663$ ).	The effect size for perceptions of human genetic variation ( $d = -.807$ ). The effect size for racial bias ( $d = -0.203$ ).	Reductions in perceptions of genetic variation were of the <b>same</b> magnitude across all three studies ( $Q = 1.72, df = 2, p = 0.424, I^2 = 0\%$ ). Reductions in racial bias were of <b>different</b> magnitudes across all three studies ( $Q = 8.64, df = 2, p = 0.013, I^2 = 76.84\%$ ).

situated within small group discussions for many days to produce a substantial reduction in racial bias. A computer-based intervention could be unsuccessful at reducing racial bias or even backfire in some schools.

Social norms surrounding the acceptability of racist speech at the schools we sampled could also be an enabling factor for the reductions in racial bias we observed. Monteiro, de França, and Rodrigues (2009) have found that older White children from working-class families in Portugal impede blatantly racist thoughts when they sense

antiracist norms in their cultural environment. But they express these racially biased thoughts when they believe it is socially acceptable. As we were only able to sample school districts with research priorities aligned with reducing stereotyping, it is possible that we only observed reductions in racial bias because the students we sampled were changing their beliefs to act in accordance with the social norms of their school culture. However, experiments conducted in US public high schools rather than Portuguese psychology laboratories have found that social norms strongly affect discriminatory behavior but have no impact on beliefs (Paluck & Shepherd, 2012). Furthermore, there are good reasons to think that the racial bias effects we observed were not due to social desirability bias.

First, we demonstrated that reductions in racial bias were mediated by changes in perceptions of human genetic variation induced by our interventions. Approximately 18–44% of the total reduction in racial bias was significantly associated with changes in perceptions of human genetic variation. Arguably, then, at least 18–44% of the changes in racial bias we observed in each RCT was due to learning and not to social desirability bias. This implies that the remaining reductions in racial bias we observed actually are attributable to social desirability bias. Yet, in the Supporting Information, we report additional findings from RCT 3 to demonstrate that the intervention effects on racial bias were not moderated by social desirability bias. Studies could estimate the impact of social desirability bias through a factorial design that randomizes a humane genetics education along with the presence or absence of clear antiracist norms. If the effects of a humane genetics education are larger when antiracist norms are present, and if these effects are not mediated by changes in how students perceive human genetic variation, then this would undercut our argument that learning about human genetic variation reduces racial bias.

A related alternative explanation for our findings is an experimenter demand effect, which is when people change their behavior in a study to conform to the subjective interpretations of the purpose of a study (Zizzo, 2008). Yet, experimenter demand effects are part and parcel of a humane genetics education. Schooling is a cultural endeavor where students become enculturated into a group by learning how to be competent users of the group's conceptual tools (Brown, Collins, & Duguid, 1989; Lave & Wenger, 1991). Our interventions educated students about the conceptual tools that geneticists have used to combat racism in the culture of science. Thus, a student's subjective interpretation of the purpose of a humane genetics education is part of the power of it. We want learners to identify with the scientists who have challenged scientific racism and we want them to apply the tools of genetics to combat prejudice in their social worlds. The ultimate purpose of a humane genetics education is to identify with the social norms of antiracist scientific culture.

Of course, our studies cannot evaluate whether our interventions achieved that ultimate purpose. Our findings say nothing about the domain transfer of these effects or their duration beyond 3 weeks in the schools we sampled. We have no idea how the knowledge learned in these interventions interacted with students' social goals, values, motivations, knowledge, and beliefs to affect their behavior in social domains outside of the biology classroom when students feel that race is salient for judging social behavior. Additionally, longitudinal studies may find that these effects wane or reverse after 3 weeks because students misremember what they learned (Lewandowsky et al., 2012). We found no evidence of backfiring at our California school site, but this does not mean that our intervention did not backfire at other sites or over longer time periods. In the Supporting Information we address the issue of intervention backfiring in more depth through a Bayesian analysis of our data. There we argue that the probability of intervention backfiring is relatively small compared with the probability of a clinically significant reduction in racial bias.

## 7 | IMPLICATIONS

In closing, the interaction between cultural context and a humane genetics education needs to be explored before generalizations about the impact of this learning on social behavior are warranted. Nevertheless, our findings provide *tentative* proof of concept of the social cognitive consequences of a humane genetics education. By helping learners construct accurate perceptions of human genetic variation we reduced racial bias. Previous studies have

found that repeated exposure to racial terminology in the biology curriculum can significantly increase belief in genetic essentialism because it leads students to perceive too much genetic variation between races (Donovan, 2017). As our study is the first to show that the opposite effect is also possible, these results tentatively suggest that biology education is a lever that affects the development of racial bias, for better or worse, when it affects how students perceive genetic variation between racially defined populations. When biology education increases the perception that races are genetically different it can increase racial bias, and when it reduces such perceptions, it can decrease racial bias. If this mechanism is correct and generalizable, then the implications of our results for biology education in racially diverse democracies could not be more apparent in an era when White nationalism is gaining political power in the United States and in Europe (Jacobs, 2015).

Racist political movements use biological essentialism to justify the oppression of minority groups (Omi & Winant, 1994). For example, discredited ideas about the biology of race were used by opponents of *Brown vs. Board of Education* in arguments to overturn it (Jackson & Depew, 2017). That opposition was halted, in part, because other scientists used population genetics to discredit the essentialist assumptions upon which opposition to school integration rested (Jackson & Depew, 2017). Ideas about human genetic variation have been and will continue to be important to policy debates about racial inequality (Byrd & Ray, 2015). Belief in genetic essentialism is still predictive of opposition to racially ameliorative policies in White (Byrd & Ray, 2015) and non-White adults in the United States (Soylu et al., 2017).

As a major purpose of science education is to help learners understand the science behind public policy debates (National Research Council, 2012; Osborne & Monk, 2000) we cannot ignore issues of race when we teach about human genetic variation. Unfortunately, current evidence suggests that when adolescents in predominantly White schools read textbooks describing the prevalence of genetic diseases in racial groups, it significantly increases their belief in genetic essentialism (Donovan, 2014, 2016, 2017). This increase reduces support for policies that redress racial inequality in education probably because it increases students' tendencies to explain racial disparities in education with genes (Donovan, 2016, 2017). If a 21<sup>st</sup> century biology education is to prevent the scientific racism of the past as we move into the genomic future, then it will need to offer youth a more humane genetics education than it currently does.

Such a proposition raises questions about the politics of official knowledge (Apple, 1993), particularly who is empowered to teach a humane genetics education? Regarding this question, it is important to point out that, depending on which country is sampled, Castéra and Clément (2014) estimate that 3–62% of biology teachers in European, South-American, African, or Middle-Eastern countries believe that, "Ethnic groups are genetically different and that is why some are superior to others" (e.g., 3% in France; 18% in Senegal; 34% in Poland; and 62% in Lebanon). Even seminomadic herders in Mongolia exhibit biological essentialist beliefs about ethnic groups (Gil-White, 2001) Essentialist thinking appears at a nonzero prevalence in every human population in which scholars have gone looking for it (Henrich, Heine, & Norenzayan, 2010). In the United States, representative studies estimate that 4% of pre-K-12 educators believe that racial inequalities are mainly due to a lower inborn potential to learn among African Americans (Quinn, 2017).

Arguably, a humane genetics education could produce inhumane outcomes in the hands of such teachers, and there could be some of these teachers anywhere that genetics is taught. However, we demonstrated in RCT 2 that genetic essentialism can be reduced, at least temporarily, in a sample of US adults through our intervention. Thus, it is possible that educating teachers about human genetic variation to challenge faulty assumptions about racial difference could reduce racial biases in the teaching population (see Larkin, 2012 for arguments supporting the plausibility of this claim; see Philip, 2011 for arguments that problematize this claim). Moreover, there are plenty of teachers who do not believe that racial inequalities are due to genetics and these educators could be the first to be empowered to teach a more humane genetics education. Science educators have begun to outline the subject matter knowledge that teachers must possess to teach about genetic variation to reduce racism (Donovan, 2015a) and they have outlined curriculum frameworks for that teaching (Beckwith et al., 2017; Donovan, 2015b; Hubbard, 2017a). These frameworks can orient the interested educator toward a more humane genetics education. However,

more research needs to be done to chart a path through what is certain to be complex educational terrain. Those interested in navigating that path should know that when science shapes our perceptions of human genetic variation it can indirectly shape our prejudices as well. Therefore, teaching and learning about human genetic variation is not socially neutral, and we should not treat it as such.

## ACKNOWLEDGMENTS

We would like to thank Okhee Lee, Andrei Cimpian, Stephany Rose Spaulding, Jessaca Spybrook, Bryan Brown, Noah Rosenberg, Ann Morning, Dan Schwartz, Carol Dweck and Jonathan Shemwell for their feedback on our work. RCT 1 was supported through the Stanford Interdisciplinary Graduate Research Fellowship. Also, this material is based upon work supported by the National Science Foundation under Grant No. 1660985. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

## CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

## ORCID

Brian M. Donovan  <http://orcid.org/0000-0003-2329-4198>

K. C. Busch  <http://orcid.org/0000-0002-4549-5401>

Susan Kowalski  <http://orcid.org/0000-0002-2012-6378>

## REFERENCES

- 23andMe Research Team, Cognitive Genomics Consortium (COGNET), Social Science Genetic Association Consortium, Lee, J. J., Wedow, R., Okbay, A., ... Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, 50(8), 1112–1121. <https://doi.org/10.1038/s41588-018-0147-3>
- Aboud, F. E., & Fenwick, V. (1999). Exploring and evaluating school-based interventions to reduce prejudice. *Journal of Social Issues*, 55(4), 767–786.
- Aikenhead, G. S. (1996). Science education: Border crossing into the subculture of science. *Studies in Science Education*, 27(1), 1–52. <https://doi.org/10.1080/03057269608560077>
- Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science*, 333, 1096–1097.
- Allport Gordon, W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Alvermann, D. E., Smith, L. C., & Readence, J. E. (1985). Prior knowledge activation and the comprehension of compatible and incompatible text. *Reading Research Quarterly*, 20(4), 420–436. <https://doi.org/10.2307/747852>
- Andreychik, M. R., & Gill, M. J. (2014). Do natural kind beliefs about social groups contribute to prejudice? Distinguishing bio-somatic essentialism from bio-behavioral essentialism, and both of these from entitativity. *Group Processes & Intergroup Relations*, 18(4), 454–474.
- Apple, M. (1993). The politics of official knowledge: Does a national curriculum make sense? *The Teachers College Record*, 95(2), 222–241.
- ASHG Denounces Attempts to Link Genetics and Racial Supremacy (2018). *The American Journal of Human Genetics*, 103(5), 636. <https://doi.org/10.1016/j.ajhg.2018.10.011>
- Baram-Tsabari, A., Sethi, R. J., Bry, L., & Yarden, A. (2006). Using questions sent to an Ask-A-Scientist site to identify children's interests in science. *Science Education*, 90(6), 1050–1072. <https://doi.org/10.1002/sce.20163>
- Baram-Tsabari, A., & Yarden, A. (2009). Identifying meta-clusters of students' interest in science and their change with age. *Journal of Research in Science Teaching*, 46(9), 999–1022. <https://doi.org/10.1002/tea.20294>
- Baram-Tsabari, A., & Yarden, A. (2010). Quantifying the gender gap in science interests. *International Journal of Science and Mathematics Education*, 9(3), 523–550. <https://doi.org/10.1007/s10763-010-9194-7>

- Bastian, B., & Haslam, N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology, 42*(2), 228–235. <https://doi.org/10.1016/j.jesp.2005.03.003>
- Beckwith, J., Bergman, K., Carson, M., Doerr, T., Geller, L., Pierce, R., ... Zicherman, C. (2017). Using dialogues to explore genetics, ancestry, and race. *The American Biology Teacher, 79*(7), 525–537.
- Beckwith, J. (2013). The persistent influence of failed scientific ideas. In S. Krimsky, & J. Gruber (Eds.), *Genetic explanations: Sense and nonsense* (368). Cambridge, MA: Harvard University Press.
- Beckwith, J. R. (2009). *Making genes, making waves: A social activist in science*. Cambridge, MA: Harvard University Press.
- Beckwith, J. R., & Pierce, R. (2018). Genes and human behavior: Ethical implications. In R. T. Gerlai (Ed.), *Molecular-genetic and statistical techniques for behavioral and neural research* (pp. 599–622). Academic Press. : Elsevier
- Becquet, C., Patterson, N., Stone, A. C., Przeworski, M., & Reich, D. (2007). Genetic structure of chimpanzee populations. *PLOS Genetics, 3*(4), e66.
- Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science, 16*(3), 162–166. <https://doi.org/10.1111/j.1467-8721.2007.00496.x>
- Boas, F. (1911). *The mind of primitive man*. New York, NY: Macmillan.
- Boas, F. (1928). *Anthropology and modern life*. New York, NY: W. W. Norton & Company.
- Bowling, B. V., Acra, E. E., Wang, L., Myers, M. F., Dean, G. E., Markle, G. C., ... Huether, C. A. (2008). Development and evaluation of a genetics literacy assessment instrument for undergraduates. *Genetics, 178*(1), 15–22. <https://doi.org/10.1534/genetics.107.079533>
- Bratsberg, B., & Rogeberg, O. (2018). Flynn effect and its reversal are both environmentally caused. *Proceedings of the National Academy of Sciences of the United States of America, 115*(26), 6674–6678. <https://doi.org/10.1073/pnas.1718793115>
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42.
- Bruckner, H., Morning, A., & Nelson, A. (2005). The expression of biological concepts of race. In *Annual Meeting of the American Sociological Association, Philadelphia, PA*.
- Bryan, L. A., & Atwater, M. M. (2002). Teacher beliefs and cultural models: A challenge for science teacher preparation programs. *Science Education, 86*(6), 821–839. <https://doi.org/10.1002/sce.10043>
- Byrd, W. C., & Ray, V. E. (2015). Ultimate attribution in the genetic era: White support for genetic explanations of racial difference and policies. *The Annals of the American Academy of Political and Social Science, 661*(1), 212–235. <https://doi.org/10.1177/0002716215587887>
- Carter, P. L. (2012). *Stubborn roots: Cultivating cultural flexibility and equity in US and south african schools*. New York, NY: Oxford University Press.
- Carter, P. L., & Reardon, S. F. (2014). *Inequality matters*. William T. Grant Foundation Paper. New York: William T. Grant Foundation. Retrieved from <https://ed.stanford.edu/sites/default/files/inequalitymatters.pdf>
- Caspi, A., & Moffitt, T. E. (2006). Gene–environment interactions in psychiatry: Joining forces with neuroscience. *Nature Reviews Neuroscience, 7*(7), 583–590.
- Castéra, J., & Clément, P. (2014). Teachers' conceptions about the genetic determinism of human behaviour: A survey in 23 countries. *Science & Education, 23*(2), 417–443.
- Chao, M. M., Hong, Y., & Chiu, C. (2013). Essentializing race: Its implications on racial categorization. *Journal of Personality and Social Psychology, 104*(4), 619–634.
- Chinn, C. A., & Malhotra, B. A. (2002). Children's responses to anomalous scientific data: How is conceptual change impeded? *Journal of Educational Psychology, 94*(2), 327–343. <https://doi.org/10.1037/0022-0663.94.2.327>
- Cimpian, A., & Salomon, E. (2014). The inheritance heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. *Behavioral and Brain Sciences, 37*(05), 461–480. <https://doi.org/10.1017/S0140525X13002197>
- Clark, W. A. V. (1991). Residential preferences and neighborhood racial segregation: A test of the Schelling segregation model. *Demography, 28*(1), 1–19. <https://doi.org/10.2307/2061333>
- Cohen, G. L., & Sherman, D. K. (2014). The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology, 65*(1), 333–371. <https://doi.org/10.1146/annurev-psych-010213-115137>
- Coop, G., Eisen, M. B., Nielsen, R., Przeworski, M., & Rosenberg, N. A. (2014). Letters: "A troublesome inheritance." Stanford Center for Computational, Evolutionary, and Human Genomics, Stanford, CA. Retrieved from <https://cehg.stanford.edu/letter-from-population-geneticists>.
- Crawford, N. G., Kelly, D. E., Hansen, M. E. B., Beltrame, M. H., Fan, S., Bowman, S. L., ... Tishkoff, S. A. (2017). Loci associated with skin pigmentation identified in African populations. *Science, 358*(6365), 1–25. <https://doi.org/10.1126/science.aan8433>
- Dar-Nimrod, I., & Heine, S. J. (2011). Genetic essentialism: On the deceptive determinism of DNA. *Psychological Bulletin, 137*(5), 800–818. <https://doi.org/10.1037/a0021860>

- Dauer, J. T., Momsen, J. L., Speth, E. B., Makohon-Moore, S. C., & Long, T. M. (2013). Analyzing change in students' gene-to-evolution models in college-level introductory biology: Analyzing change in students' gene-to-evolution models. *Journal of Research in Science Teaching*, 50(6), 639–659. <https://doi.org/10.1002/tea.21094>
- DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3), 177–188.
- Devlin, B., Daniels, M., & Roeder, K. (1997). The heritability of IQ. *Nature*, 388, 908–912.
- diSessa, A. A., & Sherin, B. L. (1998). What changes in conceptual change? *International Journal of Science Education*, 20, 1115–1191.
- Dobzhansky, T. (1937). *Genetics and the origin of species* Vol. 11). New York, NY: Columbia University Press.
- Donovan, B. M. (2014). Playing with fire? The impact of the hidden curriculum in school genetics on essentialist conceptions of race. *Journal of Research in Science Teaching*, 51(4), 462–496. <https://doi.org/10.1002/tea.21138>
- Donovan, B. M. (2015a). Putting humanity back into the teaching of human biology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 52, 65–75. <https://doi.org/10.1016/j.shpsc.2015.01.011>
- Donovan, B. M. (2015b). Reclaiming race as a topic of the U.S. biology textbook curriculum. *Science Education*, 99(6), 1092–1117. <https://doi.org/10.1002/sce.21173>
- Donovan, B. M. (2016). Framing the genetics curriculum for social justice: An experimental exploration of how the biology curriculum influences beliefs about racial difference. *Science Education*, 100(3), 586–616. <https://doi.org/10.1002/sce.21221>
- Donovan, B. M. (2017). Learned inequality: Racial labels in the biology curriculum can affect the development of racial prejudice. *Journal of Research in Science Teaching*, 54(3), 379–411. <https://doi.org/10.1002/tea.21370>
- Doron, C.-O. (2012). Race and genealogy: Buffon and the formation of the concept of "race". E. Casetta & V. Tripodi, *Making Sense of Gender, Sex, Race, and the Family, Humana. Mente Journal of Philosophical Studies*, 22, 75–109.
- Dougherty, M. J. (2009). Closing the gap: Inverting the genetics curriculum to ensure an informed public. *The American Journal of Human Genetics*, 85(1), 6–12. <https://doi.org/10.1016/j.ajhg.2009.05.010>
- Edge, M. D., & Rosenberg, N. A. (2015). Implications of the apportionment of human genetic diversity for the apportionment of human phenotypic diversity. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 52, 32–45. <https://doi.org/10.1016/j.shpsc.2014.12.005>
- Emerson, M. O., Kimbro, R. T., & Yancey, G. (2002). Contact theory extended: The effects of prior racial contact on current social ties. *Social Science Quarterly*, 83(3), 745–761.
- Emmons, N. A., & Kelemen, D. A. (2015). Young children's acceptance of within-species variation: Implications for essentialism and teaching evolution. *Journal of Experimental Child Psychology*, 139, 148–160. <https://doi.org/10.1016/j.jecp.2015.05.011>
- Engberg, M. E. (2004). Improving intergroup relations in higher education: A critical examination of the influence of educational interventions on racial bias. *Review of educational research*, 74(4), 473–524. <https://doi.org/10.3102/00346543074004473>
- Enoch, M.-A. (2012). The influence of gene–environment interactions on the development of alcoholism and drug dependence. *Current Psychiatry Reports*, 14(2), 150–158. <https://doi.org/10.1007/s11920-011-0252-9>
- Evans, E. M., Spiegel, A. N., Gram, W., Frazier, B. N., Tare, M., Thompson, S., & Diamond, J. (2010). A conceptual guide to natural history museum visitors' understanding of evolution. *Journal of Research in Science Teaching*, 47(3), 326–353. <https://doi.org/10.1002/tea.20337>
- Feldman, M. (2014). Echoes of the past: Hereditarianism and a troublesome inheritance. *PLOS Genetics*, 10(12), e1004817. <https://doi.org/10.1371/journal.pgen.1004817>
- Feldman, M., & Lewontin, R. (1975). The heritability hang-up. *Science*, 190(4220), 1163–1168. <https://doi.org/10.1126/science.1198102>
- Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist*, 54(1), 5–20.
- Ford, M. J. (2012). A dialogic account of sense-making in scientific argumentation and reasoning. *Cognition and Instruction*, 30(3), 207–245. <https://doi.org/10.1080/07370008.2012.689383>
- Foster, M. W. (2009). Looking for race in all the wrong places: Analyzing the lack of productivity in the ongoing debate about race and genetics. *Human Genetics*, 126(3), 355–362. <https://doi.org/10.1007/s00439-009-0674-1>
- Fryer, R. G., & Levitt, S. D. (2004). The causes and consequences of distinctively black names. *The Quarterly Journal of Economics*, 119(3), 767–805.
- Gardner, P. L., & Tamir, P. (1989). Interest in biology. Part I: A multidimensional construct. *Journal of Research in Science Teaching*, 26(5), 409–423.
- Gelman, S. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences*, 8(9), 404–409.
- Gil-White, F. J. (2001). Are ethnic groups biological "species" to the human brain? Essentialism in our cognition of some social categories. *Current Anthropology*, 42(4), 515–553. <https://doi.org/10.1086/321802>
- Gould, S. J. (1996). *The mismeasure of man*. New York, NY: WW Norton & Company.
- Graves, J. L. (2015). Great is their sin: Biological determinism in the age of genomics. *The Annals of the American Academy of Political and Social Science*, 661(1), 24–50. <https://doi.org/10.1177/0002716215586558>

- Gregoire, M. (2003). Is it a challenge or a threat? A dual-process model of teachers' cognition and appraisal processes during conceptual change. *Educational Psychology Review*, 15(2), 147–179.
- Guzzetti, B. J., Snyder, T. E., Glass, G. V., & Gamas, W. S. (1993). Promoting Conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly*, 28(2), 116–159. <https://doi.org/10.2307/747886>
- Haberstick, B. C., Lessem, J. M., Hewitt, J. K., Smolen, A., Hopfer, C. J., Halpern, C. T., ... Mullan Harris, K. (2014). MAOA genotype, childhood maltreatment, and their interaction in the etiology of adult antisocial behaviors. *Biological Psychiatry*, 75(1), 25–30. <https://doi.org/10.1016/j.biopsych.2013.03.028>
- Heddy, B. C., & Sinatra, G. M. (2013). Transforming misconceptions: Using transformative experience to promote positive affect and conceptual change in students learning about biological evolution. *Science Education*, 97(5), 723–744. <https://doi.org/10.1002/sce.21072>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83. <https://doi.org/10.1017/S0140525X0999152X>
- Hong, J.-L., Shim, K.-C., & Chang, N.-K. (1998). A study of Korean middle school students' interests in biology and their implications for biology education. *International Journal of Science Education*, 20(8), 989–999.
- Hubbard, A. R. (2017a). Teaching race (bioculturally) matters: A visual approach for college biology courses. *The American Biology Teacher*, 79(7), 516–524. <https://doi.org/10.1525/abt.2017.79.7.516>
- Hubbard, A. R. (2017b). Testing common misconceptions about the nature of human racial variation. *The American Biology Teacher*, 79(7), 538–543. <https://doi.org/10.1525/abt.2017.79.7.538>
- Hudson, N. (1996). From “nation” to “race”: The origin of racial classification in eighteenth-century thought. *Eighteenth-Century Studies*, 29(3), 247–264.
- Hunley, K. L., Cabana, G. S., & Long, J. C. (2016). The apportionment of human diversity revisited. *American Journal of Physical Anthropology*, 160(4), 561–569. <https://doi.org/10.1002/ajpa.22899>
- Jablonski, N. G., & Chaplin, G. (2010). Human skin pigmentation as an adaptation to UV radiation. *Proceedings of the National Academy of Sciences of the United States of America*, 107(Suppl. 2), 8962–8968.
- Jackson, J. P., Jr., & Depew, D. J. (2017). *Darwinism, democracy, and race: American Anthropology and evolutionary biology in the twentieth century*. New York, NY: Routledge.
- Jacobs, T. S. (2015). *White pride world wide: A contemporary history of online, international right-wing extremist ideology and culture*. University of Oslo. Retrieved from <https://www.duo.uio.no/handle/10852/45238>
- Jamieson, A., & Radick, G. (2013). Putting Mendel in his place: How curriculum reform in genetics and counterfactual history of science can work together. In K. Kampourakis (Ed.), *The Philosophy of biology* Vol. 1, pp. 577–595). Dordrecht: Springer. Retrieved from [http://link.springer.com/10.1007/978-94-007-6537-5\\_25](http://link.springer.com/10.1007/978-94-007-6537-5_25).
- Jayarathne, T. E., Ybarra, O., Sheldon, J. P., Brown, T. N., Feldbaum, M., Pfeffer, C. A., & Petty, E. M. (2006). White Americans' genetic lay theories of race differences and sexual orientation: Their relationship with prejudice toward blacks, and gay men and lesbians. *Group Processes & Intergroup Relations*, 9(1), 77–94. <https://doi.org/10.1177/1368430206059863>
- Jensen, A. (1969). How much can we boost IQ and scholastic achievement. *Harvard Educational Review*, 39(1), 1–123.
- Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, 25(6), 881–919.
- Kahan, B. C., & Morris, T. P. (2013). Assessing potential sources of clustering in individually randomised trials. *BMC Medical Research Methodology*, 13(58), 1–9.
- Kaplan, J. M., & Winther, R. G. (2013a). Prisoners of abstraction? The theory and measure of genetic variation, and the very concept of “race”. *Biological Theory*, 7(4), 401–412.
- Kaplan, J. M., & Winther, R. G. (2013b). *Realism, antirealism, and conventionalism about “race”*. Retrieved from <http://philpapers.org/rec/KAPRAA>
- Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology*, 88(4), 686–702. <https://doi.org/10.1037/0022-3514.88.4.686>
- Keller, M. C. (2014). Gene×environment interaction studies have not properly controlled for potential confounders: The problem and the (simple) solution. *Biological Psychiatry*, 75(1), 18–24. <https://doi.org/10.1016/j.biopsych.2013.09.006>
- Kimel, S. Y., Huesmann, R., Kunst, J. R., & Halperin, E. (2016). Living in a genetic world how learning about interethnic genetic similarities and differences affects peace and conflict. *Personality and Social Psychology Bulletin*, 42(5), 688–700.
- Kontopantelis, E., & Reeves, D. (2016). metaan: Random-effects meta-analysis. In T. M. Palmer, & J. A. C. Sterne (Eds.), *Meta-analysis in Stata: An updated collection from the Stata Journal* (2nd ed., pp. 55–67). College Station, TX: STATA press.
- Larkin, D. (2012). Using the conceptual change model of learning as an analytic tool in researching teacher preparation for student diversity. *Teachers College Record*, 114(8), 1–35.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press. Retrieved from [http://books.google.com/books?hl=en&lr=&id=CAVIOwW3vYAC&oi=fnd&pg=PA11&dq=lave+1991&ots=OAqzqwQFFk&sig=a50\\_7z3isT6rYY3MdfDsFo5nwx](http://books.google.com/books?hl=en&lr=&id=CAVIOwW3vYAC&oi=fnd&pg=PA11&dq=lave+1991&ots=OAqzqwQFFk&sig=a50_7z3isT6rYY3MdfDsFo5nwx).

- Lehrer, R., & Schauble, L. (2004). Modeling natural variation through distribution. *American Educational Research Journal*, 41(3), 635–679.
- Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, 347(6219), 262–265.
- Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology*, 74(6), 1421–1436.
- Lewontin, R. (1996). *Biology as ideology: The doctrine of DNA*. Toronto, Ontario: House of Anansi.
- Lewontin, R. C. (1972). The apportionment of human diversity. *Evolutionary Biology*, 6, 381–398.
- Livingstone, F. B., & Dobzhansky, T. (1962). On the non-existence of human races. *Current Anthropology*, 3(3), 279–281.
- Lynch, J., Bevan, J., Achter, P., Harris, T., & Condit, C. M. (2008). A preliminary study of how multiple exposures to messages about genetics impact on lay attitudes towards racial and genetic discrimination. *New Genetics and Society*, 27(1), 43–56. <https://doi.org/10.1080/14636770701843634>
- Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3), 106–131. <https://doi.org/10.1177/1529100612451018>
- MacMahon, B. (1968). Gene-environment interaction in human disease. *Journal of Psychiatric Research*, 6, 393–402.
- Mallon, R. (2013). Was Race thinking invented in the modern West? *Studies in History and Philosophy of Science Part A*, 44(1), 77–88. <https://doi.org/10.1016/j.shpsa.2012.10.001>
- Markus, H., & Moya, P. M. L. (2011). *Doing race: 21 essays for the 21st century*. New York, NY: WW Norton & Company.
- Monteiro, M. B., de França, D. X., & Rodrigues, R. (2009). The development of intergroup bias in childhood: How social norms can shape children's racial behaviours. *International Journal of Psychology*, 44(1), 29–39. <https://doi.org/10.1080/00207590802057910>
- Morin-Chassé, A., Suhay, E., & Jayaratne, T. E. (2017). Discord over DNA: Ideological responses to scientific communication about genes and race. *The Journal of Race, Ethnicity, and Politics*, 2, 1–40. <https://doi.org/10.1017/rep.2017.17>
- Morning, A. (2009). Toward a sociology of racial conceptualization for the 21<sup>st</sup> century. *Social Forces*, 87(3), 1167–1192.
- Morning, A. J. (2011). *The nature of race: How Scientists think and teach about human difference*. Berkeley, CA: University of California Press.
- Morton, T. A., Hornsey, M. J., & Postmes, T. (2009). Shifting ground: The variable use of essentialism in contexts of inclusion and exclusion. *British Journal of Social Psychology*, 48(1), 35–59. <https://doi.org/10.1348/014466607X270287>
- National Research Council (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.
- Next Generation Science Standards (2016). *Standards by topic*. Retrieved from <http://www.nextgenscience.org/overview-topics>
- Okbay, A., Beauchamp, J. P., Fontana, M. A., Lee, J. J., Pers, T. H., Rietveld, C. A., ... Benjamin, D. J. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature*, 533, 1–4. <https://doi.org/10.1038/nature17671>
- Omi, M., & Winant, H. (1994). *Racial formation in the United States: From the 1960s to the 1990s*. New York, NY: Psychology Press.
- Opfer, J. E., Nehm, R. H., & Ha, M. (2012). Cognitive foundations for science assessment design: Knowing what students know about evolution. *Journal of Research in Science Teaching*, 49(6), 744–777. <https://doi.org/10.1002/tea.21028>
- Osborne, J., & Monk, M. (2000). *Good practice in science teaching: What research has to say*. Berkshire: Open University Press.
- Osborne, J., Donovan, B. M., Henderson, J. B., MacPherson, A. C., & Wild, A. (2016). *Arguing from evidence in middle school science: 24 activities for productive talk and deeper learning*. Thousand Oaks, CA: Corwin Press. Retrieved from <https://us.corwin.com/en-us/nam/arguing-from-evidence-in-middle-school-science/book250184#description>.
- Paluck, E. L., & Shepherd, H. (2012). The salience of social referents: A field experiment on collective norms and harassment behavior in a school social network. *Journal of Personality and Social Psychology*, 103(6), 899–915. <https://doi.org/10.1037/a0030015>
- Parrott, R. L., Silk, K. J., Dillow, M. R., Krieger, J. L., Harris, T. M., & Condit, C. M. (2005). Development and validation of tools to assess genetic discrimination and genetically based racism. *Journal of the National Medical Association*, 97(7), 980–990.
- Pauker, K., Ambady, N., & Apfelbaum, E. P. (2010). Race salience and essentialist thinking in racial stereotype development. *Child Development*, 81(6), 1799–1813.
- Pauker, K., Xu, Y., Williams, A., & Biddle, A. M. (n.d.). Race essentialism and social contextual differences in children's racial stereotyping. *Child Development*. Advance online publication. Retrieved from [http://manoa.hawaii.edu/isplab/wp-content/uploads/2015/10/Pauker-Xu-Williams-Biddle-in-press\\_ChildDevelopment.pdf](http://manoa.hawaii.edu/isplab/wp-content/uploads/2015/10/Pauker-Xu-Williams-Biddle-in-press_ChildDevelopment.pdf).
- Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, 49(1), 65–85.
- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90(5), 751–783.

- Philip, T. M. (2011). An "ideology in pieces" approach to studying change in teachers' sensemaking about race, racism, and racial justice. *Cognition and Instruction*, 29(3), 297–329.
- Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research*, 63(2), 167–199.
- Putnam, R. D. (2007). E pluribus unum: Diversity and community in the twenty-first century the 2006 Johan Skytte Prize Lecture. *Scandinavian Political Studies*, 30(2), 137–174.
- Quinn, D. M. (2017). Racial attitudes of preK–12 and postsecondary educators: Descriptive evidence from nationally representative data. *Educational Researcher*, 46, 397–411. 0013189X17727270. <https://doi.org/10.3102/0013189X17727270>
- Reich, D. (2018). *Who we are and how we got here: Ancient DNA and the new science of the human past*. Oxford University Press.
- Relethford, J. H. (1997). Hemispheric difference in human skin color. *American Journal of Physical Anthropology*, 104(4), 449–457.
- Relethford, J. H. (2002). Apportionment of global human genetic diversity based on craniometrics and skin color. *American Journal of Physical Anthropology*, 118(4), 393–398.
- Rosenberg, N. A. (2011). A population-genetic perspective on the similarities and differences among worldwide human populations. *Human Biology*, 83(6), 659–684. <https://doi.org/10.3378/027.083.0601>
- Rosenberg, N. A., & Edge, M. D. (in press). Genetic clusters and the race debates: a perspective from population genetics. In Q. Spencer (Ed.), *The race debates from metaphysics to medicine*. Oxford, UK: Oxford University Press.
- Rosenberg, N. A., Edge, M. D., Pritchard, J. K., & Feldman, M. W. (2018). Interpreting polygenic scores, polygenic adaptation, and human phenotypic differences. *Evolution, Medicine, and Public Health*, 2019(1), 26–34. <https://doi.org/10.1093/emph/eoy036>
- Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. A., & Feldman, M. W. (2002). Genetic structure of human populations. *Science*, 298(5602), 2381–2385.
- Ryoo, K., & Linn, M. C. (2012). Can dynamic visualizations improve middle school students' understanding of energy in photosynthesis? *Journal of Research in Science Teaching*, 49(2), 218–243.
- Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, 16(4), 475–522.
- Shemwell, J., Chase, C. C., & Schwartz, D. L. (2015). Seeking the general explanation: A Test of inductive activities for learning and transfer. *Journal of Research in Science Teaching*, 52(1), 58–83.
- Shtulman, A., & Schulz, L. (2008). The relation between essentialist beliefs and evolutionary reasoning. *Cognitive Science: A Multidisciplinary Journal*, 32(6), 1049–1062. <https://doi.org/10.1080/03640210801897864>
- Sinatra, G. M., Brem, S. K., & Evans, E. M. (2008). Changing minds? Implications of conceptual change for teaching and learning about biological evolution. *Evolution: Education and Outreach*, 1(2), 189–195. <https://doi.org/10.1007/s12052-008-0037-8>
- Sinatra, G. M., Kienhues, D., & Hofer, B. K. (2014). Addressing challenges to public understanding of science: Epistemic cognition, motivated reasoning, and conceptual change. *Educational Psychologist*, 49(2), 123–138. <https://doi.org/10.1080/00461520.2014.916216>
- Sinatra, G. M., Southerland, S. A., McConaughy, F., & Demastes, J. W. (2003). Intentions and beliefs in students' understanding and acceptance of biological evolution. *Journal of Research in Science Teaching*, 40(5), 510–528. <https://doi.org/10.1002/tea.10087>
- Smedley, A., & Smedley, B. D. (2005). Race as biology is fiction, racism as a social problem is real: Anthropological and historical perspectives on the social construction of race. *American Psychologist*, 60(1), 16–26.
- Solomon, R. L., & Lessac, M. S. (1968). A control group design for experimental studies of developmental processes. *Psychological Bulletin*, 70(3), 145–150.
- Soylu, Y. N., Estrada-Villalta, S., & Adams, G. (2017). The (biological or cultural) essence of essentialism: Implications for policy support among dominant and subordinated groups. *Frontiers in Psychology*, 8, 900. <https://doi.org/10.3389/fpsyg.2017.00900>
- Stern, F., & Kampourakis, K. (2017). Teaching for genetics literacy in the post-genomic era. *Studies in Science Education*, 53(2), 193–225. <https://doi.org/10.1080/03057267.2017.1392731>
- Stewart, J., Cartier, J. L., & Passmore, C. M. (2004). Developing understanding through model-based inquiry. In S. M. Donovan, & J. D. Bransford (Eds.), *How students learn: Science in the classroom* (pp. 515–563). Washington, DC: National Academies Press.
- Storage, D., Horne, Z., Cimpian, A., & Leslie, S.-J. (2016). The frequency of "brilliant" and "genius" in teaching evaluations predicts the representation of women and African Americans across fields. *PLOS One*, 11(3), e0150194.
- Strike, K. A., & Posner, G. J. (1992). A revisionist theory of conceptual change. In R. A. Duschl, & R. J. Hamilton (Eds.), *Philosophy of science, cognitive psychology, and educational theory and practice* (pp. 147–176). Albany, NY: State University of New York Press. Retrieved from [https://books.google.com/books?hl=en&lr=&id=bJpW2WvfkC&oi=fnd&pg=PA147&dq=conceptual+change+strike+and+posner+1992&ots=1derTn4Azf&sig=EOck7eOdFcsrD9yshE6b\\_hoVbHU](https://books.google.com/books?hl=en&lr=&id=bJpW2WvfkC&oi=fnd&pg=PA147&dq=conceptual+change+strike+and+posner+1992&ots=1derTn4Azf&sig=EOck7eOdFcsrD9yshE6b_hoVbHU).
- Taie, S., & Goldring, R. (2018). *Characteristics of public elementary and secondary school teachers in the United States: Results from the 2015–16 National Teacher and Principal Survey First Look (NCES 2017-072rev)*. Washington, DC: US Department of Education, National Center for Educational Statistics. p. 50

- Tawa, J. (2016). Belief in race as biological: Early life influences, intergroup outcomes, and the process of “unlearning”. *Race and Social Problems*, 8(3), 244–255. <https://doi.org/10.1007/s12552-016-9176-7>
- The Daily Stormer . (2016, September 4). *A Normie’s Guide to the Alt-Right*. Retrieved from <https://katana17.wordpress.com/2016/09/04/the-daily-stormer-a-normies-guide-to-the-alt-right/>
- Turkheimer, E., Haley, A., Waldron, M., D’onofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, 14(6), 623–628.
- Uitto, A. (2014). Interest, attitudes and self-efficacy beliefs explaining upper-secondary school students’ orientation towards biology-related careers. *International Journal of Science and Mathematics Education*, 12(6), 1425–1444. <https://doi.org/10.1007/s10763-014-9516-2>
- Van den Broek, P. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science*, 328(5977), 453–456.
- Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.
- Wade, N. (2014). *A troublesome inheritance: Genes, Race and Human History*. Melbourne: Penguin Press.
- Weiss, K. M. (2008). The good, the bad, and the ugly. *Evolutionary Anthropology: Issues, News, and Reviews*, 17(3), 129–134. <https://doi.org/10.1002/evan.20163>
- Weiss, M. J., Lockwood, J. R., & McCaffrey, D. F. (2016). Estimating the standard error of the impact estimator in individually randomized trials with clustering. *Journal of Research on Educational Effectiveness*, 9(3), 421–444. <https://doi.org/10.1080/19345747.2015.1086911>
- White Pride For Kids. (n.d.). Retrieved from <https://www.stormfront.org/kids/>
- Wilkerson, M. H., & Laina, V. (2017). Youth reasoning with interactive data visualizations: A preliminary study. In P. Blikstein & D. Abrahamson (Eds.), *Proceedings of the 2017 Conference on Interaction Design and Children* (pp. 411–416). New York, NY: ACM.
- Wilkerson-Jerde, M. H., Gravel, B. E., & Macrander, C. A. (2015). Exploring shifts in middle school learners’ modeling activity while generating drawings, animations, and computational simulations of molecular diffusion. *Journal of Science Education and Technology*, 24(2–3), 396–415.
- Williams, M. J., & Eberhardt, J. L. (2008). Biological conceptions of race and the motivation to cross racial boundaries. *Journal of Personality and Social Psychology*, 94(6), 1033–1047.
- Willinsky, J. (1998). *Learning to divide the world: Education at empire’s end*. Minneapolis, MN: University of Minnesota Press.
- Wolf, E. R., Kahn, J. S., Roseberry, W., & Wallerstein, I. (1994). Perilous ideas: Race, culture, people [and comments and reply]. *Current Anthropology*, 35(1), 1–12.
- Yudell, M., Roberts, D., DeSalle, R., & Tishkoff, S. (2016). Taking race out of human genetics. *Science*, 351(6273), 564–565.
- Yzerbyt, V., Corneille, O., & Estrada, C. (2001). The interplay of subjective essentialism and entitativity in the formation of stereotypes. *Personality and Social Psychology Review*, 5(2), 141–155.
- Zizzo, D. J. (2008). Experimenter demand effects in economic experiments. Available at SSRN: <https://ssrn.com/abstract=1163863> or <http://dx.doi.org/10.2139/ssrn.1163863>.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Donovan BM, Semmens R, Keck P, et al. Toward a more humane genetics education: Learning about the social and quantitative complexities of human genetic variation research could reduce racial bias in adolescent and adult populations. *Science Education*. 2019;1–32.

<https://doi.org/10.1002/sce.21506>